La implementación como disciplina fundamental de la ciencia de datos empresariales

Thomas Davenport

Katie Malone

***Nota de los editores de la columna*:** *En este artículo, nos centramos en un problema clave en el sector: lograr que los modelos de ciencia de datos se implementen en la producción dentro de las organizaciones. Las tareas y habilidades que conlleva la implementación no se consideran normalmente un componente clave de las iniciativas de ciencia de datos, pero son fundamentales para el éxito de la misma. Describimos las pruebas del problema de implementación, los componentes de esta y cómo algunos programas presenciales de la licenciatura en Analítica Empresarial intentan enseñar habilidades de implementación.*

**Palabras clave***:* ciencia de datos, implementación, análisis empresarial, gestión del cambio

Cada vez parece más evidente que la implementación –lograr que los sistemas analíticos y de inteligencia artificial (IA) se utilicen plena y exitosamente en las organizaciones– se está convirtiendo en una de las disciplinas fundamentales en todas las fases de un proyecto de ciencia de datos empresariales. A menudo pensamos erróneamente que, cuando el modelo o el sistema basado en algoritmos se pone en producción como parte de un proceso empresarial, la implementación de un modelo analítico o de ciencia de datos es la última etapa del proceso. Un aspecto clave de la implementación es el cambio en el proceso empresarial: un modelo implementado con éxito tomará un conjunto de tareas que antes eran manuales, basadas en la heurística o simplemente imposibles, e introducirá una solución basada en algoritmos. Empezar primero con el algoritmo y solo pensar, hasta el final del proyecto, en cómo introducirlo en el proceso empresarial es donde muchas implementaciones fracasan. En lugar de pensar en la implementación como el último paso de un conjunto lineal de actividades, un científico de datos –o, al menos, los miembros clave de los equipos de ciencia de datos– debería considerar los factores que tienen un fuerte efecto en la implementación a lo largo del proyecto de ciencia de datos.

Por ejemplo, en el proyecto a gran escala Orion, de UPS, a través del cual se pretendía crear rutas optimizadas en tiempo real para los choferes de la empresa, la mayor parte de los diez años que duró el ciclo de implementación y un presupuesto de varios cientos de millones de dólares se fueron en problemas de implementación y gestión del cambio (Centro de Tecnología y Sostenibilidad, 2016). El complejo algoritmo se tenía que implementar en servidores y dispositivos portátiles, convencer a los choferes y a sus supervisores de que el algoritmo funcionaba eficazmente y modificar múltiples procesos de clasificación y carga de paquetes. El de UPS no es un caso único: muchas empresas con adaptaciones exitosas impulsadas por datos han tenido que superar obstáculos de implementación similares para obtener beneficios. Sin embargo, las habilidades de implementación, como el análisis de partes interesadas y la gestión del cambio, no reciben –comparativamente– la misma atención en los programas tradicionales de formación en ciencia de datos, lo cual deja a muchos científicos de datos sin la preparación adecuada para esta difícil parte del proceso.

El reto que representa la implementación exige que los científicos de datos se anticipen a los problemas de implementación y sean capaces de responder a ellos. Sin embargo, muchos de ellos no están capacitados para identificar y resolver los problemas de implementación, por lo cual podrían pensar que la implementación no es parte de su trabajo. Anecdóticamente, al preguntarles a los científicos de datos o a los líderes en el campo de la ciencia de datos qué porcentaje de sus modelos ha sido implementado, pocos de ellos lo saben (Davenport, 2019).

Algunas organizaciones dejan que los científicos de datos se centren en la creación de modelos y en la codificación mientras otras funciones se encargan de las cuestiones organizativas o técnicas de la implementación. En Southern California Edison, por ejemplo, un científico de datos trabaja de manera conjunta con un «consultor de analítica predictiva», el cual se encarga de las relaciones con el Departamento de Clientes y es responsable de las cuestiones de cambio organizativo (Davenport, 2020b). Estas funciones también se denominan «gestores de productos», «traductores de analítica» o «estrategas de datos/analítica/IA» (Henke *et al.*, 2018). Otras empresas hacen hincapié en las funciones de «ingeniería de datos» o «ingeniería de aprendizaje automático» al gestionar las cuestiones técnicas de la implementación. Un importante banco con el que trabajamos se refería a estas funciones como «la versión descafeinada» de un analista cuantitativo, una caracterización que puede ser técnicamente correcta, pero que menosprecia la importancia que tienen en el éxito de los proyectos.

Cada vez se reconoce más que los científicos de datos no pueden ser «unicornios» y poseer todos los tipos de conocimientos necesarios sobre proyectos (Davenport, 2020a). Por lo tanto, a menudo es conveniente asignar las principales responsabilidades de implementación a alguna persona del equipo de ciencia de datos que esté capacitada y que será evaluada en función del éxito que tenga la implementación general del proyecto. Sin embargo, como mínimo, incluso aquellos científicos de datos más inclinados a los modelos y a la codificación deben ser conscientes de los problemas de implementación y señalarlos a los especialistas en caso de que existan. Si los modelos no se implementan, tanto estos como los científicos de datos que los crean podrían terminar siendo de escasa utilidad para los empleadores. Todo el ámbito de la ciencia de datos podría perder terreno dentro de una organización si los modelos solo se implementan en contadas ocasiones. Y, en el caso de aquellos sectores en los que la auditabilidad y la transparencia son absolutamente cruciales, como la banca, las finanzas y el cuidado de la salud, un modelo mal implementado supone un riesgo legal, empresarial o sanitario.

# Evidencias de un problema de implementación

Encuestas realizadas a organizaciones e informes de investigación de mercado elaborados en Estados Unidos y en todo el mundo sugieren que los retos de la implementación se han generalizado. Las encuestas realizadas por Rexer Analytics a científicos de datos a lo largo de varios años revelaron que solo una pequeña parte de los encuestados afirmó que todos sus modelos se habían implementado. Karl Rexer, quien realiza las encuestas, comentó: «La implementación de modelos predictivos sigue siendo un reto importante para muchas empresas... En nuestra Encuesta de Ciencia de Datos de 2017, solo el 13 % de los científicos de datos afirma que sus modelos siempre se implementan. Y la implementación no tienen visos de mejora: en cada encuesta, remontándonos hasta 2009, cuando hicimos esta pregunta por primera vez, vemos que los resultados son casi idénticos» (Alteryx, 2017). En una de las encuestas, las empresas en las que era más probable que se implementaran modelos tenían científicos de datos con un nivel de satisfacción laboral sustancialmente más alto (Allen *et al.*, 2015).

Una encuesta realizada en 2019 por NewVantage Partners a grandes empresas de servicios financieros y ciencias biológicas reveló que las empresas estaban adoptando activamente las tecnologías y soluciones de IA, y el 91.5 % de las empresas informaron que estaban haciendo inversiones continuas en IA. Pero solo el 14.6 % de las empresas informaron que habían implementado funciones de IA en la producción generalizada (NewVantage Partners, 2019).

En una encuesta global de McKinsey de 2019, cuyo título era «La adopción de la IA demuestra su utilidad, pero pocos dimensionan su impacto», entre el 12 % (en bienes de consumo envasados) y el 54 % (en empresas de alta tecnología) de los encuestados tenía al menos una aplicación de aprendizaje automático implementada en un proceso o producto. Solo el 30 % de los encuestados en general utilizaba la IA en productos o procesos en múltiples unidades de negocio y funciones (Cam *et al.*, 2019).

En una [encuesta realizada](https://www.bcg.com/publications/2020/is-your-company-embracing-full-potential-of-artificial-intelligence) en 2020 por [MIT Sloan Management Review/BCG](https://www.bcg.com/publications/2020/is-your-company-embracing-full-potential-of-artificial-intelligence) a ejecutivos de todo el mundo, solo una de cada diez empresas afirmó haber obtenido beneficios financieros significativos al implementar la IA en sus empresas (Ransbotham *et al.*, 2020). Los autores del informe de la encuesta destacan la importancia de la implementación: «Pero es la fase final de madurez de la IA, esto es, orquestar con éxito las macro y microinteracciones entre humanos y máquinas lo que realmente permite sacarle partido».

Las empresas de investigación de mercado también han llegado a la conclusión de que muchas organizaciones se enfrentan a difíciles problemas de implementación. Por ejemplo, en un informe de Gartner sobre el futuro de la IA, se comentaba que «la realidad es que la mayoría de las organizaciones tienen dificultades para convertir los proyectos piloto de IA en una producción a escala empresarial, lo que limita la capacidad de aprovechar el potencial comercial de la IA» (Costello, 2020).

Forrester, otro investigador de mercado, sugiere que es necesario un amplio abanico de habilidades para lograr una exitosa implementación de la IA en la empresa (no solo en el aspecto técnico):

La IA es, por naturaleza, un sistema de aprendizaje, lo cual hace necesario un ciclo de desarrollo continuo para potenciar su contribución empresarial. Dada la necesidad de gestionar el cambio empresarial, los proyectos de IA requieren una amplia experiencia que va más allá del desarrollo y la implementación de tecnología. Generar confianza en las decisiones automatizadas es esencial para llegar a una IA que sea aceptada por clientes y empleados. (Granzen, 2020)

Estas fuentes respaldan la idea de que la implementación es un aspecto importante, pero poco planteado, de la ciencia de datos. Sin embargo, para abordar la cuestión de la implementación, necesitamos comprender mejor qué es.

# ¿Cuáles son los componentes de la implementación?

Un primer paso fundamental es entender el problema empresarial y el entorno en el que se aplicará el modelo. Esta comprensión requiere que el científico de datos observe en detalle el proceso actual y a las personas que lo realizan. Incluye evaluar el proceso actual de toma de decisiones, entender qué datos podrían estar disponibles y cómo se recopilan (si se recopilan entre personas o máquinas) y determinar las habilidades y capacidades de las personas que hacen el trabajo hoy en día (Birnbaum, 2004; Horrell *et al.*, 2020). Incorporar el resultado a un producto implica entender las necesidades del cliente y la manera en que los productos existentes las satisfacen o no. En el caso de los productos tradicionales, este suele ser el ámbito de un jefe de producto. Incluso, cuando los equipos de ciencia de datos cuentan también con una función de gestión de productos, el hecho de que el experto técnico (es decir, el científico de datos) esté en contacto directo con los usuarios es importante para establecer el contexto de la forma en que se utilizará finalmente el modelo.

En esta etapa inicial de la ciencia de datos, el científico de datos adquiere información fundamental sobre los tipos de errores que puede cometer el futuro modelo y las implicaciones de cada uno de ellos, lo que le permitirá ajustar el modelo durante su desarrollo. Esta es también la etapa en la que un científico de datos reflexivo empezará a pensar qué aspectos del modelo deberán monitorearse cuando se implemente, y cómo mostrar esas métricas a las partes interesadas adecuadas.

Otro precursor importante de la implementación es ganarse la confianza del director o de otras partes interesadas responsables del proceso o del producto. Dichas personas tomarán la decisión de implementar o no el nuevo modelo y es posible que los detalles técnicos de su funcionamiento nunca sean accesibles para los interesados. Por lo tanto, es importante que confíen en que el científico de datos entiende el problema y sabrá cómo resolverlo, y que el modelo no introducirá nuevas complicaciones en el ámbito empresarial.

Antes de construir el modelo, el científico de datos también debe estar familiarizado con la información de la pila tecnológica en la que encajará el modelo. Un modelo altamente implementable estará estrechamente integrado en los sistemas existentes, no requerirá muchas habilidades o tareas nuevas por parte del usuario ni ningún cambio importante en la arquitectura tecnológica. Aquellos que requieran cambios tecnológicos importantes, naturalmente, llevarán más tiempo, costarán más y serán más difíciles de implementar. Cuando los modelos requieren una fuerte adaptación técnica para su implementación, el científico de datos debe ampliar su fase de investigación y creación de confianza para incluir al personal de TI.

En este punto, la «infraestructura» también incluye las fuentes de datos y las estructuras a través de las cuales los datos fluyen de su fuente al algoritmo. Uno de los fallos de los modelos que parecían funcionar perfectamente durante su desarrollo es que, al llegar el momento de implementarlo, los datos que se introducen en los algoritmos no son fiables, no se actualizan con regularidad, cambian de forma imprevisible o presentan alguna diferencia con los que se utilizaron para desarrollar el modelo. Alguien del equipo de analítica o del proyecto de IA debe asegurarse de que los datos sean accesibles, de suficiente calidad y en volúmenes lo suficientemente grandes como para apoyar el desarrollo y la implementación del modelo. También es necesario confiar en que los datos o el modelo no contengan sesgos y tengan un propósito ético. El «zar de datos de IA» es una función que carece de un título consensuado y que muchas organizaciones olvidan, por lo que usualmente se incluye en la función de científico de datos (Beck *et al.*, 2019).

En la medida en que los problemas de implementación puedan ser supervisados y gestionados con tecnología, esta debe ser considerada (algunas herramientas y plataformas de ciencia de datos, como Domino Data Lab, DataRobot, Databricks y H20.ai tienen funciones que dan soporte o automatizan aspectos de la implementación, las cuales vale la pena considerar en esta etapa, incluso si no eran parte del proceso de desarrollo del modelo). Para los muchos problemas que un científico de datos no puede anticipar en esta etapa, ya que los sistemas de datos de producción tienden a encontrar nuevas e interesantes formas de romperse todo el tiempo, tener a un humano revisando el modelo es por lo general una buena idea. En particular, los importantes conocimientos técnicos necesarios para este aspecto de la implementación conducen a otra división del trabajo; la función de «ingeniero de aprendizaje automático» se emplea cada vez más en este punto del proceso.

A medida que se construye y perfecciona el modelo, el científico de datos tiene que ocuparse no solo de la adecuación del modelo a los datos, sino también de la responsabilidad por los cambios que genera en el resultado del proceso, tanto de forma inmediata como en el futuro. Las decisiones que el científico de datos tome en esta etapa pueden tener un impacto significativo en la facilidad con que se harán esas futuras actualizaciones. Si un modelo complejo y no explicable se ajusta a los datos mejor que uno explicable, ¿vale la pena que el mayor ajuste del modelo reduzca la probabilidad de implementación? Si las nuevas funciones requieren nuevos datos, ¿podrán proporcionarlos los usuarios de primera línea? Si el modelo requiere un procesador basado en la GPU para correr, ¿funcionará con la arquitectura de los sistemas existentes?

Por último, una vez desarrollado el modelo y creado un programa (o una interfaz de programa de aplicación -API-) para la implementación de la solución aún quedarán tareas por realizar. Un científico de datos o un miembro del equipo centrado en la implementación debería monitorear el rendimiento del modelo y de los sistemas relacionados cuando se pongan en marcha por primera vez. ¿Los usuarios emplean el modelo y el sistema de forma adecuada y toman las decisiones correctas? ¿Los primeros casos aportan valor económico? ¿Ofrece el modelo resultados coherentes y reproducibles, especialmente a medida que se actualiza el conjunto de datos subyacente? ¿Las predicciones del modelo presentan algún problema ético imprevisto o muestran un sesgo? Algún miembro del equipo de ciencia de datos debería desempeñar en este punto la función de economista, proporcionando información sobre si el nuevo modelo está logrando el retorno de la inversión. Sin embargo, muchos científicos de datos que no han sido formados en esta fase de «modo de mantenimiento» del desarrollo de modelos suelen pasar a construir el siguiente modelo llegado este punto, dejando un sistema altamente técnico y a menudo difícil de depurar para que otro se ocupe de él.

Por lo tanto, debería ser evidente que muchos de estos factores que impulsan una implementación exitosa no se consideran o se enseñan típicamente como ciencia de datos. Sin embargo, si estos factores no se gestionan bien en un proceso de implementación de ciencia de datos, es probable que no se obtenga ningún valor general. Cualquier proceso que permita a los científicos de datos o a los analistas llevar a cabo un mejor y más relevante análisis de datos debería ser una parte integral de la ciencia de datos. Las organizaciones que no cuentan con la infraestructura de ciencia de datos (no solo técnica, sino también organizativa y estratégica) que permitiría a un científico de datos especializarse en la construcción de modelos deberían crear dicha infraestructura o incorporar la implementación en la descripción del puesto de trabajo de los científicos de datos. Si no hacen ninguna de las dos cosas, están condenando a los científicos de datos centrados en modelos a tener muchas dificultades para causar algún impacto.

# Enseñanza y desarrollo de habilidades de implementación

¿Qué opciones hay para asegurarse de que los científicos de datos están bien preparados para implementar modelos en el trabajo? La educación tradicional ha mostrado (hasta ahora) algunas deficiencias, dejando que empleadores y científicos de datos dependan en gran medida de sí mismos. Pero, a medida que el campo madure, los educadores expertos que sean conscientes de que existe una «brecha de implementación» podrán abordarla en los planes de estudio.

El sistema de educación superior estadounidense suele tener bastante éxito en la enseñanza de muchas de las otras habilidades que se esperan de los científicos de datos, como la programación informática, la estadística y el aprendizaje automático. Sin embargo, dado que muchas de las capacidades de implementación son contextuales y están basadas en relaciones, es difícil enseñarlas en un curso tradicional dirigido por un instructor. La única forma segura de desarrollarlas es crear experiencias en las que se implemente un proyecto de ciencia de datos. Incluso allí puede ser difícil establecer la responsabilidad de los modelos, las decisiones y las acciones desarrolladas en los cursos o proyectos finales.

Los programas de posgrado en ciencia de datos se enfrentan a retos para incorporar estas experiencias de implementación. Dichos programas suelen encontrarse en escuelas de informática e ingeniería, las cuales quizás no tengan relación con empresas que patrocinen el trabajo en proyectos, y no suelen incluir el cambio organizativo u otros temas de implementación en sus planes de estudio. Las escuelas de negocios suelen tener esas relaciones y suelen exigir una o dos clases de comportamiento organizativo y cambio, pero algunos profesores tal vez no estén convencidos de la necesidad de proyectos prácticos si van en detrimento de más instrucción técnica. Además, muchos programas de maestría en ciencia de datos son exclusivamente *online*, un medio que no se presta a muchas experiencias en la vida real. Los programas de doctorado que conducen a las carreras de Ciencias de Datos –desde Física, hasta Estadística, pasando por Antropología– no suelen centrarse en cuestiones de implementación empresarial, ni siquiera en un solo curso.

Para la enseñanza de los aspectos técnicos de la implementación, una opción prometedora son los programas de maestría en Analítica Empresarial que se imparten de manera presencial, pues es más probable que incluyan contenidos orientados a la implementación y un proyecto en el que se traten cuestiones sobre la misma. Uno de los primeros cursos de este tipo (que comenzó en 2007), la [maestría en Analítica](file:///C%3A%5CUsers%5Cmariasilviavidiri%5CDesktop%5CTRABAJOS%20A%20REALIZAR%5CVetting%5Cmaestri%CC%81a%20en%20Anali%CC%81tica) (https://analytics.ncsu.edu/) de la Universidad Estatal de Carolina del Norte, cuenta con un proyecto de prácticas en equipo con empresas, agencias gubernamentales y organizaciones no lucrativas que dura ocho meses. Michael Rappa, director del curso, comentó en un correo electrónico (11 de junio de 2020) lo siguiente:

Los estudiantes aprenden que el aspecto clave del trabajo de un científico de datos es un proceso: entender la necesidad empresarial, idear un marco analítico, manejar los datos, construir modelos, extraer información, comunicar los resultados y, si estos lo justifican, empaquetar un producto que pueda implementarse en un entorno de producción. Todo tiene que ver con el proceso. Si el equipo mantiene la integridad del proceso, el resultado final será útil, aunque no se utilice. Si el equipo elude el proceso o socava su integridad, entonces lo que entregue no servirá de nada: por muy prometedor que parezca, no se puede confiar en los resultados. Los estudiantes aprenden a respetar el rigor del proceso enfrentándose a las muchas formas en que puede arruinarse cuando no se es disciplinado. Esto puede ser frustrante –y aleccionador– para aquellos estudiantes habituados a los problemas de los libros de texto. Las lecciones aprendidas se les quedarán grabadas.

Para entender cómo encaja la implementación en el proceso de modelado de principio a fin, la realización de proyectos en colaboración con el sector proporciona a los estudiantes una valiosa experiencia en el mundo real. Por ejemplo, en el MIT, el relativamente nuevo programa [Maestría en Analítica Empresarial](https://mitsloan.mit.edu/master-of-business-analytics) cuenta con un laboratorio de analítica con proyectos empresariales, un proyecto final de siete meses y un curso que trata las habilidades relacionadas con la implementación. Los tres cursos se describen a continuación:

Laboratorio de Analítica: Vincula a equipos de estudiantes con proyectos de vanguardia relacionados con la analítica, el aprendizaje automático o las tecnologías digitales en su aplicación a cuestiones y problemas empresariales.

Proyecto Final: Este proyecto de siete meses asigna a dos estudiantes a organizaciones que proponen un proyecto de analítica y normalmente incluye tanto el trabajo a distancia durante un semestre, como una residencia de verano de diez semanas en el lugar del proyecto.

De la Analítica a la Acción: Desarrolla el aprecio por la dinámica organizativa y la competencia para navegar por las redes sociales, trabajar en equipo, desmitificar las recompensas y los incentivos, comprender las iniciativas de cambio y tomar decisiones acertadas.

Michelle Li, directora del Programa de Analítica Empresarial, escribió en un correo electrónico (22 de junio de 2020) sobre el propósito de estos cursos:

En estos cursos, los estudiantes deben trabajar en equipo en problemas analíticos prácticos con organizaciones asociadas. Al hacerlo, los estudiantes tienen que enfrentarse no solo a conjuntos de datos desordenados, dispares e incompletos, sino también a barreras organizativas y culturales que pueden dificultar la aplicación de las técnicas de modelización académica en el mundo real. Una vez que los estudiantes de la maestría concluyen el programa de inmersión de un año, están mejor preparados para resolver problemas de ciencia de datos de gran envergadura que requieren habilidades tanto técnicas como orientadas a la implementación, como la comunicación empresarial, la comprensión del comportamiento organizacional y un enfoque en la creación de relaciones.

Los científicos de datos, al sopesar sus opciones académicas, deben evaluar la experiencia en proyectos que obtendrán en cada institución, del mismo modo en que considerarían el renombre de la escuela y los planes de estudio.

Naturalmente, existen formas en que los estudiantes pueden adquirir conocimientos y experiencia relacionados con la implementación fuera de un programa educativo formal. A menos que el trabajo en proyectos y la experiencia de implementación se conviertan en la norma de la educación en ciencias de datos, los empleadores y científicos de datos deberán asumir que gran parte de la responsabilidad de enseñar y aprender este importante conjunto de habilidades recae sobre ellos. Los trabajos anteriores, por ejemplo, pueden haber proporcionado a un determinado científico de datos lecciones en la práctica sobre la implementación. Incluso, si la experiencia laboral no incluye la ciencia de datos o la analítica, otros tipos de proyectos ciertamente pueden fortalecer las habilidades y sensibilidades orientadas a la implementación. Una sugerencia para los gerentes de ciencia de datos es que hagan que los científicos de datos nóveles, recién contratados, empiecen con una «rotación», manteniendo o ampliando un modelo en producción, como forma de servicio a la organización y como entrenamiento. Poder ver el estado final desde una etapa temprana puede generar un estado de ánimo muy diferente al de empezar leyendo artículos y creando modelos, algo más habitual en las nuevas contrataciones de ciencia de datos.

# Pasos organizativos para incrementar el éxito de la implementación

Hay varios pasos que las organizaciones pueden dar para incrementar las probabilidades de que los proyectos se implementen con éxito. Entre ellos se encuentran:

* Crear equipos de ciencia de datos cuyos miembros se especialicen en ciertas habilidades, pero que todos contribuyan en la implementación. Los equipos pueden incluir científicos de datos orientados a la creación de modelos, ingenieros de aprendizaje automático, gestores de productos, ingenieros de datos y miembros de operaciones técnicas, quienes, conjuntamente, serán responsables de la implementación, y cada uno de ellos contribuirá a la misma de alguna manera.
* Al contratar científicos de datos y miembros del equipo de ciencia de datos, indagar sobre su experiencia con implementaciones durante las entrevistas de trabajo. Cierto director de Analítica que conocemos da por terminadas abruptamente las entrevistas si el candidato no logra recordar y describir alguna implementación exitosa en trabajos anteriores.
* Dejar claras las responsabilidades relacionadas con la implementación que tendrán los científicos de datos y los miembros del equipo de ciencia de datos en las descripciones de los puestos de trabajo y las asignaciones de proyectos.
* Establecer un «canal» hacia la implementación para cada proyecto de ciencia de datos que culmine en la implementación de producción. Al igual que ocurre con los procesos de desarrollo de productos en «fase inicial» o con las actividades de desarrollo de productos *lean*, que comienzan con productos con una viabilidad mínima, puede haber desgaste en el camino, pero el objetivo es lograr la implementación plena. Farmers Insurance es una empresa que ha establecido este tipo de estructura (Davenport y Bean, 2018). Y, en una encuesta de Deloitte de 2018, el 54 % de los encuestados, ejecutivos estadounidenses en grandes empresas que trabajan con IA, afirmó que ya contaban con un proceso para pasar los prototipos a la producción (Deloitte Insights, 2018).
* Crear una «cobertura aérea» para el éxito de las implementaciones promoviendo la aceptación de los proyectos entre los miembros de la alta dirección. En una gran empresa farmacéutica, por ejemplo, el director de Datos y Analítica obtuvo la aprobación por consenso de la alta dirección para tres grandes proyectos, cada uno con un valor considerable para la empresa. El compromiso incluía suficientes recursos humanos y financieros para lograr la plena implementación de los proyectos, suponiendo que se cumplieran los objetivos intermedios.
* Si una organización tiene una cultura orientada a los datos –caracterizada por buenas prácticas en cuanto a la incentivación del uso de datos, el fomento de la alfabetización de datos entre los empleados y el apoyo a la toma de decisiones basada en datos por parte de los directivos–, una iniciativa de ciencia de datos tendrá un terreno mucho más fértil para lograr su adopción.

Naturalmente, algunos científicos de datos empleados en empresas podrían simplemente no estar interesados en adquirir habilidades orientadas a la implementación, pues quizás sientan que su única función es crear algoritmos eficaces e innovadores. Pero, a final de cuentas, la gran mayoría de los trabajos para científicos de datos empresariales implican la resolución de problemas empresariales con datos y análisis, y el algoritmo es tan solo un medio para ese fin. Sin embargo, si solo se orientan a la creación de algoritmos, deben ser conscientes de que un modelo o algoritmo que no se implementa –por muy brillante que sea técnicamente– significa que el trabajo no se ha terminado y que sus organizaciones sacarán escaso provecho de sus esfuerzos. Como mínimo, los científicos de datos deben respetar la necesidad que hay de habilidades orientadas a la implementación y a las personas que las aportan a los proyectos.

# Agradecimientos

Agradecemos los comentarios de tres revisores anónimos y del editor en jefe de *HDSR*.

# Declaración de divulgación

Thomas Davenport es asesor de tres empresas mencionadas en esta columna: NewVantage Partners, DataRobot y Deloitte. También es investigador en el MIT, pero no está vinculado a la maestría en Analítica Empresarial de dicho instituto.

# Bibliografía

Allen, H., Gearan, P. & Rexer, K. (2015). «2015 data science survey». *Rexer Analytics*. [www2.cs.uh.edu/~ceick/UDM/Rexer2015.pdf](http://www2.cs.uh.edu/~ceick/UDM/Rexer2015.pdf)

*Alteryx*. (2017, 12 de septiembre). «Alteryx Promote announced: New offering to easily deploy, manage and integrate data science models for real-time decisions» [Press release]. <https://www.alteryx.com/press-releases/2017-09-12-alteryx-promote-announced-new-offering-easily-deploy-manage-and-integrate>

Beck, M., Davenport, T. & Libert, B. (2019, 14 de marzo). «The AI roles some companies forget to fill». *Harvard Business Review*. <https://www.hbr.org/2019/03/the-ai-roles-some-companies-forget-to-fill>

Birnbaum, M. (2004). «Human research and data collection via the Internet». *Annual Review of Psychology*, *55*, 803–832*.* <https://www.doi.org/10.1146/annurev.psych.55.090902.141601>

Cam, A., Chui, M. & Hall, B. (2019, 22 de noviembre). «Global AI Survey: AI proves its worth, but few scale impact». *McKinsey & Company.* <https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact>

Center for Technology and Sustainability. (2016, 30 de marzo).«Looking under the hood: ORION technology adoption at UPS». *Business for Social Responsibility.* <https://www.bsr.org/en/our-insights/case-study-view/center-for-technology-and-sustainability-orion-technology-ups>

Costello, K. (2020, 5 de febrero). «Gartner predicts the future of AI technologies». *Gartner*. <https://www.gartner.com/smarterwithgartner/gartner-predicts-the-future-of-ai-technologies/>

Davenport, T. (2019, 21 de marzo). «What’s your deployment score? » *International Institute for Analytics.* <https://www.iianalytics.com/blog/2019/3/21/whats-your-deployment-score>

Davenport, T. (2020a, 19 de mayo). «Beyond unicorns: Educating, classifying, and certifying business data scientists». *Harvard Data Science Review*, *2*(2).<https://doi.org/10.1162/99608f92.55546b4a>

Davenport, T. (2020b, 30 de julio). «Machine learning and organizational change at Southern California Edison». *Forbes.* https://www.[forbes.com/sites/tomdavenport/2020/07/30/machine-learning-and-organizational-change-at-southern-california-edison/#1767a96f3336](https://www.forbes.com/sites/tomdavenport/2020/07/30/machine-learning-and-organizational-change-at-southern-california-edison/#1767a96f3336)

Davenport, T. & Bean, R. (2018,1 de agosto). «Farmers accelerates its time to impact with AI». *Forbes*. <https://www.forbes.com/sites/tomdavenport/2018/08/01/farmers-accelerates-its-time-to-impact-with-ai/?sh=56308d4b672a>

*Deloitte Insights*. (2018). «State of AI in the enterprise» (2nd ed.). <https://www2.deloitte.com/content/dam/insights/us/articles/4780_State-of-AI-in-the-enterprise/DI_State-of-AI-in-the-enterprise-2nd-ed.pdf>

Granzen, A. (2020, 17 de febrero). «Consultancies are reinventing their service model for AI». *Forrester*. <https://go.forrester.com/blogs/consultancies-are-reinventing-their-service-model-for-ai/>

Henke, N., Levine, J. & McInerney, P. (2018, 5 de febrero). «You don’t have to be a data scientist to fill this must-have analytics role». *Harvard Business Review*. <https://hbr.org/2018/02/you-dont-have-to-be-a-data-scientist-to-fill-this-must-have-analytics-role>

Horrell, M., McElhinney, A. & Reynolds, L. (2020, 30 de abril). «Data science in heavy industry and the Internet of Things». *Harvard Data Science Review 2*(2). <https://doi.org/10.1162/99608f92.834c6595>

*NewVantage Partners*. (2019). «Big Data and AI Executive Survey of 2019». [https://www.newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf](https://newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf)

Ransbotham, S., Khodabandeh, S., Kiron, D., Candelon, F., Chu, M. & Lafountain, B. (2020, 19 de octubre). «Expanding AI’s impact with organizational learning». *MIT Sloan Management Review*. <https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/?utm_medium=pr&utm_source=release&utm_campaign=ReportBCGAI2020>

©2021 Thomas Davenport y Katie Malone. Este artículo está sujeto a una [licencia internacional](https://creativecommons.org/licenses/by/4.0/legalcode) de Creative Commons Attribution (CC BY 4.0), salvo que se indique lo contrario con respecto al material concreto incluido en el artículo.

*Nota:* Este documento es una versión traducida al español del original: *Deployment as a Critical Business Data Science Discipline* de Thomas Davenport y Katie Malone, realizada bajo los términos de la licencia Creative Commons Attribution (CC BY 4.0).