

Científico de datos: *El trabajo más popular del siglo XXI*

Conozca a las personas que pueden convertir un montón de datos no estructurados y desordenados en un verdadero tesoro.

Por Thomas H. Davenport y D.J. Patil

En junio de 2006, cuando Jonathan Goldman comenzó a trabajar en LinkedIn , el sitio de redes empresariales, el lugar aún tenía el aspecto de una start-up. La empresa tenía unos 8 millones de cuentas y el número crecía rápidamente, ya que los miembros existentes invitaban a sus amigos y colegas a unirse. Sin embargo, los usuarios no buscaban conexiones con las personas que ya participaban en el sitio al nivel que lo habían previsto los ejecutivos. Al parecer, faltaba algo en la experiencia social. Como lo explicó un administrador de LinkedIn, "era como llegar a la recepción de una conferencia y darse cuenta de que son todos desconocidos. Entonces, uno se queda parado en un rincón bebiendo un trago y lo más probable es que se vaya temprano".



Goldman, que es doctor en física de Stanford, estaba intrigado por las vinculaciones que observaba y por la diversidad de los perfiles de usuario. Todo esto generaba datos desordenados y análisis difíciles de manejar, pero a medida que comenzó a explorar las conexiones de las personas, empezó a ver posibilidades. Comenzó a formular teorías, poner a prueba sus instintos y descubrir patrones que le permitieron predecir a las redes de qué personas se les sugeriría un perfil determinado. Podía imaginar que las nuevas funciones que aprovechaban la heurística que él desarrollaba podrían ofrecerles valor a los usuarios. Sin embargo, el equipo de ingeniería de LinkedIn, que estaba concentrado en los desafíos de ampliar el sitio, parecía indiferente. Algunos colegas despreciaban abiertamente las ideas de Goldman. ¿Por qué los usuarios necesitarían que LinkedIn arme sus redes en lugar de hacerlo ellos? El sitio ya contaba con un importador de libretas de direcciones que podía extraer todas las conexiones de un miembro.

Afortunadamente, Reid Hoffman, el cofundador y director general de LinkedIn en ese momento (ahora director ejecutivo), tenía fe en la potencia del análisis por sus experiencias en PayPal y le había concedido a Goldman un alto grado de autonomía. En primer lugar, le había brindado a Goldman una forma de evitar el ciclo de lanzamiento de productos tradicional mediante la publicación de pequeños módulos en forma de anuncios en las páginas más populares del sitio.

Por medio de uno de estos módulos, Goldman comenzó a probar qué sucedería si a los usuarios se les presentaran

creación de redes, como el "cierre de triángulos", la noción de que si una persona conoce a Manuel y a Susana, es factible que Manuel y Susana se conozcan entre sí. Goldman y su equipo también lograron que se pudiera responder una sugerencia con un solo clic.

Los gerentes directivos de LinkedIn no tardaron en reconocer esta excelente idea y convertirla en una función estándar. En ese momento, el negocio despegó por completo. Los anuncios de "Gente que podrías conocer" alcanzaron una tasa de clics 30% más alta que la obtenida por otros mensajes que invitaban a visitar más páginas del sitio. Generaron millones de nuevas vistas de páginas. Gracias a esta sola función, la trayectoria de crecimiento de LinkedIn se elevó de forma significativa.

Una nueva estirpe

Goldman es un buen ejemplo de los nuevos miembros clave de las organizaciones: el "científico de datos". Es un profesional de alto nivel con la capacitación y la curiosidad para hacer descubrimientos en el mundo big data. Este cargo comenzó a existir hace unos pocos años. (El nombre fue acuñado en 2008 por uno de nosotros, D.J. Patil, y Jeff Hammerbache, en ese momento, líderes de las iniciativas de datos y análisis de LinkedIn y Facebook, respectivamente). Sin embargo, ya hay miles de científicos de datos que están trabajando en start-ups y en empresas consolidadas. Su aparición repentina en la escena empresarial refleja el hecho de que las empresas ahora deben lidiar con información que viene en variedades y volúmenes nunca antes vistos. Si su organización almacena varios petabytes de datos, si la información más importante para su empresa reside en otros formatos que no sean filas y columnas de números, o si responder su mayor pregunta supondría combinar varios esfuerzos de análisis, tiene una oportunidad de big data.

Gran parte del entusiasmo actual por big data está relacionado con las tecnologías que permiten dominarlos, como Hadoop (el marco más utilizado para el procesamiento del sistema de archivos distribuidos) y las herramientas de código abierto relacionadas, la computación en la nube y la visualización de datos. Si bien esos son avances importantes, también tienen la misma importancia las personas que cuentan con las habilidades (y el enfoque) necesarias para poder aprovecharlos. En este frente, la demanda superó ampliamente la oferta. De hecho, en algunos sectores, la escasez de científicos de datos se está convirtiendo en una limitación grave. Greylock Partners, una empresa de capital de riesgo de fase temprana que respaldó a empresas como Facebook, LinkedIn, Palo Alto Networks y Workday, está tan preocupada por la escasa fuerza laboral disponible que desarrolló su propio equipo de reclutamiento especializado para atraer profesionales talentosos a las empresas que forman parte de su cartera de clientes.

En algunos sectores, la escasez de científicos de datos se está convirtiendo en una limitación grave.

personas con las que aún no se habían conectado, pero que parecían conocidas, por ejemplo, personas con las que habían compartido cargos en la escuela o en algún trabajo. Para ello, creó un anuncio personalizado que mostraba las tres mejores coincidencias nuevas para cada usuario en función de los antecedentes ingresados en su perfil de LinkedIn. En unos pocos días, ya era evidente que había sucedido algo asombroso. La tasa de clics de esos anuncios era la más alta que habían visto. Goldman continuó optimizando la generación de sugerencias e incorporó ideas para la

Una idea en pocas palabras

Hay una nueva función que se vuelve cada vez más prominente en las organizaciones: la del científico de datos. Los científicos de datos son las personas que saben cómo encontrar las respuestas a las preguntas empresariales importantes que surgen del actual tsunami de información no estructurada. A medida que las empresas se desesperan por capitalizar el potencial de los big data, la limitación más grande que muchas enfrentan es la escasez de estos profesionales especiales.

Aún no se diseñaron programas universitarios para formar científicos de datos, por lo que el reclutamiento de estos profesionales requiere creatividad. Busque personas exitosas de cualquier campo con un enfoque sólido en la informática y los datos, lo cual podría alejarlo tanto de los negocios como para considerar ámbitos como la física experimental y la biología de sistemas.

Reconozca también que las características de un trabajo que puede atraer y retener a un científico de datos

pueden diferir con respecto a los factores que atraen al resto de los profesionales. Los científicos de datos necesitan autonomía, pero quieren estar "en el puente de mando" y abordar los problemas de administración con colegas directivos en tiempo real. El dinero cuenta como una señal de valor, pero en una disciplina dinámica, la posibilidad de dejar una huella personal al trabajar con los problemas más intrigantes y aprovechar los flujos de datos más diversos tal vez tenga más peso.

"Una vez que tienen datos", afirma Dan Portillo, que dirige ese equipo, "realmente necesitan tener personas que puedan administrarlos y encontrar estadísticas en ellos".

¿Quiénes son estas personas?

Si la capitalización de big data depende de contratar a los pocos científicos de datos que hay disponibles, entonces el desafío de los administradores es aprender a identificar a esos profesionales, atraerlos a sus empresas y alcanzar una mayor productividad. Ninguna de esas tareas es tan clara como con otras funciones organizativas establecidas. Para empezar, piense que no hay programas universitarios que ofrezcan títulos en ciencia de datos. Además, tampoco hay consenso sobre en qué parte de la organización encaja la función, de qué forma los científicos de datos pueden agregar mayor valor y de qué modo se debe medir el rendimiento de esta función.

El primer paso para satisfacer la necesidad de científicos de datos, por lo tanto, consiste en comprender qué hacen en las empresas. Luego, debe preguntarse qué habilidades necesitan y en qué campos se encuentran más fácilmente esas habilidades.

Más que nada, la tarea de los científicos de datos es hacer descubrimientos mientras navegan por los datos. Es su método predilecto para navegar por el mundo que los rodea. Dado que se sienten cómodos en el terreno digital, pueden darles estructura a grandes cantidades de datos sin formato y facilitar el análisis. Identifican fuentes de datos diversas, se unen a ellas con otras fuentes potencialmente incompletas y limpian el conjunto resultante. En un plano competitivo en el que los desafíos evolucionan y los datos nunca dejan de fluir, los científicos de datos permiten que los encargados de tomar las decisiones pasen de un análisis específico para cada caso a una conversación continua con los datos.

Los científicos de datos saben que enfrentan limitaciones técnicas, pero no permiten que eso dificulte la búsqueda de soluciones novedosas. A medida que hacen descubrimientos, comunican sus hallazgos y sugieren sus implicaciones para las nuevas direcciones empresariales. Por lo general, son creativos a la hora de exponer visualmente la información y muestran los patrones encontrados de forma clara y atrac-

tiva. Aconsejan a los ejecutivos y gerentes de producto sobre las implicaciones que tienen los datos en los productos, los procesos y las decisiones.

Dado el estado emergente de su mercado, los científicos de datos suelen tener que diseñar sus propias herramientas e incluso realizar investigaciones de estilo académico. Yahoo, una de las empresas que empleó un grupo de científicos de datos de forma temprana, fue clave en el desarrollo de Hadoop. El equipo de datos de Facebook creó el lenguaje Hive para programar proyectos de Hadoop. Muchos otros científicos de datos, especialmente en empresas basadas en datos, como Google, Amazon, Microsoft, Walmart, eBay, LinkedIn y Twitter, ampliaron y perfeccionaron el kit de herramientas.

¿Qué clase de persona hace todo esto? ¿Qué capacidades necesita tener un científico de datos para ser exitoso? Piense en este profesional como un híbrido entre un pirata de datos, un analista, un comunicador y un asesor de confianza. La combinación es extremadamente potente y exclusiva.

La habilidad universal más básica de los científicos de datos es la capacidad para escribir códigos. Es posible que no lo siga siendo en los próximos cinco años, ya que muchas personas más tendrán el título de "científico de datos". Más perdurable será la necesidad de contar con científicos de datos para comunicarse en un lenguaje que comprendan todas las partes interesadas y para demostrar las habilidades especiales que se necesitan en la narración con datos, ya sea verbal, visual o, idealmente, una combinación de estas dos.

Aun así, podríamos decir que la característica predominante entre los científicos de datos es una curiosidad intensa, el deseo de sumergirse más allá de la superficie de un problema, encontrar las respuestas ocultas y sintetizarlas en un conjunto claro de hipótesis que se pueden poner a prueba. Esto, por lo general, conlleva el pensamiento asociativo que caracteriza a la mayoría de los científicos creativos de cualquier campo. Por ejemplo, sabemos de un científico de datos que, mientras estudiaba un problema de fraude, notó que este era similar a un problema sobre el tipo de segmentación del ADN. Al combinar esos dos terrenos tan dispares, él y su equipo pudieron diseñar una solución que redujo drásticamente las pérdidas por fraude. Tal vez esté comenzando a

verse con claridad por qué la palabra "científico" se ajusta a esta función emergente. Los físicos experimentales, por ejemplo, también tienen que diseñar equipos, recopilar datos, realizar diversos experimentos y comunicar sus resultados. Por ende, las empresas que buscaban personas que pudieran trabajar con datos complejos han tenido buena suerte al reclutar profesionales con formación y antecedentes laborales en las ciencias físicas y sociales. Algunos de los científicos de datos más eficaces y brillantes tienen doctorados en campos esotéricos, como ecología y biología de sistemas. George Roumeliotis, el líder de un equipo de ciencia de datos de Intuit, en Silicon Valley, tiene un doctorado en astrofísica. Lo que resulta menos sorprendente

es que muchos de los científicos de datos que actualmente trabajan en empresas antes se capacitaron en informática, matemática o economía. Pueden emerger de cualquier campo que tenga un enfoque sólido en informática y datos.

Es importante tener en mente esa imagen del científico, ya que la palabra "datos" podría desviar fácilmente la búsqueda de profesionales. Como nos indicó Portillo, "los antecedentes tradicionales de hace unos 10 o 15 años no alcanzan en la actualidad". Un analista cuantitativo puede ser muy bueno para analizar datos, pero no para manejar una gran cantidad de datos no estructurados y darles una forma específica que permita su análisis. Un experto en administración de datos puede ser muy bueno para generar y organizar datos de forma estructurada, pero no para convertir datos no estructurados en estructurados o para analizar realmente los datos. Si bien las personas sin habilidades sociales sólidas pueden progresar en profesiones de datos tradicionales, los científicos de datos deben contar con esas habilidades para ser eficaces.

Roumeliotis fue claro y nos comentó que no contrata al personal en función de capacidades estadísticas o analíticas, sino que para comenzar con la búsqueda de científicos de datos, les pregunta a los aspirantes si saben desarrollar prototipos en un lenguaje de programación convencional, como Java. Roumeliotis busca un conjunto de habilidades (conocimientos sólidos de matemática, estadística, informática y probabilidad) y ciertas actitudes. Quiere contratar a personas con talento para los problemas empresariales y empatía hacia los clientes. Luego, según afirma, a todo eso le suma capacitación en servicio y algún curso ocasional en una tecnología en particular.

Varias universidades planean lanzar programas de ciencia de datos, y algunos programas de análisis existentes, como la Maestría en Ciencias del Análisis de la Universidad Estatal de Carolina del Norte, agregan constantemente tareas y ejercicios de big data. Además, algunas empresas intentan desarrollar a sus propios científicos de datos. Luego de adquirir la empresa de big data Green-plum, EMC notó que la disponibilidad de los científicos de datos sería un factor determinante en su propia explotación de big data y en la de sus clientes. Por lo tanto, su división de Servicios Educativos lanzó un programa de capacitación y certificación en ciencia de datos y análisis de big data. EMC pone el programa a disposición de los empleados y los clientes, y algunos de sus graduados ya trabajan en iniciativas internas de big data.

A medida que proliferan las ofertas educativas, se debe expandir la canalización de profesionales. Los proveedores de tecnologías de big data también se están esforzando por que sus productos sean más fáciles de usar. Mientras tanto, a un científico de datos se le ocurrió un enfoque creativo para cerrar esta brecha. El Programa de becarios en ciencia de datos de Insight, una beca posdoctoral diseñada por Jake Klamka (con formación de físico

Cómo encontrar a los científicos de datos que necesita

1 Centre el reclutamiento en las universidades más prestigiosas en este aspecto (Stanford, MIT, Berkeley, Harvard y Carnegie Mellon) y también en otras que hayan demostrado formar grandes líderes: la Universidad Estatal de Carolina del Norte, UC Santa Cruz, la Universidad de Maryland, la Universidad de Washington y UT Austin.

2 Analice la lista de miembros de los grupos de usuarios que se dedican a las herramientas de ciencia de datos. Los grupos de usuarios de R (para la herramienta estadística de código abierto preferida por los científicos de datos) y los grupos de interés de Python (para PIGgies) son buenos puntos de partida.

3 Busque científicos de datos en LinkedIn. Casi todos figuran allí, por lo que podrá ver si cuentan con las habilidades que necesita.

4 Converse con científicos de datos en las conferencias de Strata, Structure:Data y Hadoop World, y en eventos similares (ahora hay casi uno por semana), o bien en "encuentros" informales de científicos de datos en el Área de la Bahía, Boston, Nueva York, Washington D.C., Londres, Singapur y Sídney.

5 Conozca a un inversor de capital de riesgo local, quien seguramente habrá recibido diversas propuestas de macrodatos durante el último año.

6 Realice una competición en Kaggle o TopCoder, los sitios de competición de análisis y codificación. Manténgase en contacto con los participantes más creativos.

7 No pierda el tiempo con los aspirantes que no sepan codificar. Las habilidades de codificación no deben ser de primer nivel, pero deben ser mínimamente aceptables. Además, busque evidencia de que los aspirantes aprenden rápidamente sobre los nuevos métodos y tecnologías.

8 Asegúrese de que los aspirantes comprendan la estructura de un conjunto de datos y de que puedan brindar una narrativa coherente sobre una estadística de datos clave. Pruebe si los aspirantes pueden comunicar con números, de forma visual y de forma verbal.

9 Esté atento a si los aspirantes están demasiado desconectados del mundo empresarial. Cuando les pregunta cómo su trabajo puede aplicarse a los desafíos de administración, ¿no encuentran respuesta?

10 Pregúntele a los aspirantes sobre sus estadísticas y análisis favoritos, y sobre cómo perfeccionan sus habilidades. ¿Tienen un certificado en el curso avanzado en línea de Aprendizaje automático de Stanford, colaboraron en proyectos de código abierto o crearon un repositorio de código en línea para compartir (por ejemplo, en GitHub)?

Los científicos de datos quieren crear y no solo aconsejar. Un científico describe la tarea del asesor como "la zona muerta".

de partículas), toma científicos del ámbito académico y, en seis semanas, los prepara para que se conviertan en científicos de datos exitosos. El programa combina el asesoramiento de expertos en datos de empresas locales (como Facebook, Twitter, Google y LinkedIn) con la exposición a los desafíos reales de big data. Klamka, que en un principio pensaba trabajar con 10 becarios, terminó aceptando a 30 de una base que superaba los 200 aspirantes. Ahora hay más organizaciones que esperan participar. "La demanda de las empresas ha sido asombrosa", nos comentó Klamka. "Simplemente no consiguen esta clase de profesionales calificados".

¿Por qué un científico de datos querría trabajar aquí?

Incluso a medida que crecen los rangos de los científicos de datos, la competencia para conseguir a los profesionales más capacitados sigue siendo feroz. Los candidatos evalúan las oportunidades laborales en función de cuán interesantes son los desafíos de big data. Como comentó uno de ellos, "si quisieramos trabajar con datos estructurados, estaríamos en Wall Street". Dado que actualmente los clientes potenciales más calificados no tienen un pasado empresarial, los gerentes de contratación necesitan repensar la manera en que pueden mostrar de forma atractiva el potencial de innovación que ofrecen sus problemas.

La paga, evidentemente, también será un factor importante. Un buen científico de datos tendrá muchas oportunidades laborales, por lo que sus ofertas salariales irán en aumento. Varios científicos de datos que trabajan en start-ups indicaron que habían exigido y obtenido grandes paquetes con opciones de compra de acciones. Incluso si alguien acepta un puesto por otros motivos, la remuneración refleja cierto nivel de respeto y el valor que se espera que esa función le garantice a la empresa. Sin embargo, nuestra encuesta informal sobre las prioridades de los científicos de datos reveló algo mucho más importante. Quieren estar "en el puente de mando". Esta frase hace referencia al programa de televisión de la década de 1960 *Viaje a las estrellas*, en el cual el capitán de la nave, James Kirk, dependía en gran medida de los datos proporcionados por el Dr. Spock. Los científicos de datos quieren estar al frente del desarrollo y conocer en tiempo real el dinámico conjunto de opciones que presenta.

Si se tiene en cuenta la dificultad de encontrar y retener científicos de datos, uno pensaría que una buena estrategia sería contratarlos como asesores. La mayoría de las empresas de consultoría aún no tienen planteles completos con asesores de este tipo. Incluso las empresas de mayor envergadura, como Accenture, Deloitte e IBM Global Services, están en las primeras etapas de dirigir proyectos de big data para sus clientes. Las habilidades de los científicos de datos con los que cuentan principalmente se aplican a problemas de análisis cuantitativo más convencionales. Las empresas de servicios de análisis en el extranjero, como Mu Sigma, podrían ser las que hagan las primeras incursiones importantes con científicos de datos.

Sin embargo, los científicos de datos con los que hablamos afirman que quieren crear contenido y no solo aconsejar a los encargados de tomar las decisiones. Un científico describió la tarea del asesor como "la zona muerta: todo lo que se hace es decirle a otro lo que los análisis sugieren que haga". Al crear soluciones funcionales, pueden tener un mayor impacto y dejar su huella como pioneros en la profesión.

Cuidado y atenciones

Los científicos de datos no dan buenos resultados cuando se los limita demasiado. Deben tener la libertad de experimentar y explorar posibilidades. Aun así, necesitan entablar relaciones cercanas con el resto de la empresa. Los vínculos más importantes que deben forjar son con ejecutivos a cargo de productos y servicios, y no con personas que supervisan funciones empresariales. Tal como lo demuestra la historia de Jonathan Goldman, la mayor oportunidad de agregar valor no está en la creación de informes o presentaciones para los ejecutivos superiores, sino en la innovación con los procesos y productos destinados a los clientes.

LinkedIn no es la única empresa que contrata científicos de datos con el fin de generar ideas para productos, funciones y servicios que agreguen valor. En Intuit, a los científicos de datos se les solicita que desarrollen estadísticas para consumidores y clientes de pequeñas empresas, y que envíen informes a un nuevo vicepresidente ejecutivo en relación con los big data, el diseño social y el marketing. GE ya utiliza la ciencia de datos para optimizar los contratos de servicio y los intervalos de mantenimiento de los productos industriales. Google, por supuesto, cuenta con científicos de datos para perfeccionar sus principales algo-



ritmos de búsqueda y publicidad. Zynga contrata científicos de datos para optimizar la experiencia de juego y, así, también son complicados de retener. Simplemente no hay obtener ingresos y fidelidad a largo plazo. Netflix creó muchas personas con esa combinación de conocimientos el reconocido Premio Netflix, que se otorga al equipo de científicos y habilidades informáticas y analíticas.

ciencia de datos que desarrolle el método más eficaz para mejorar el sistema de recomendaciones de películas de la empresa. La empresa de preparación de pruebas Kaplan utiliza científicos de datos para descubrir estrategias de aprendizaje eficaces.

No obstante, el hecho de que estas personas con habilidades sofisticadas en un campo dinámico trabajen entre

contratar y, dado el mercado competitivo de sus servicios, a los analistas cuantitativos de Wall Street de las décadas de 1980 y 1990. En esos días, las personas con conocimientos de física y matemática se volcaban a los bancos de inversión y los fondos de cobertura, donde podían diseñar estrategias de

En este momento, los científicos de datos son similares a los analistas cuantitativos de Wall Street de las décadas de 1980 y 1990. En esos días, las personas con conocimientos de física y matemática se volcaban a los bancos de inversión y los fondos de cobertura, donde podían diseñar estrategias de datos y algoritmos completamente nuevos. Luego, diversas universidades desarrollaron programas de maestría en ingeniería financiera, los cuales forjaron una segunda generación de profesionales que era más accesible para las empresas convencionales. Más tarde, en la década de 1990, se repitió el patrón con los ingenieros de búsqueda, cuyas habilidades intrincadas pronto comenzaron a enseñarse en los programas de informática.

La pregunta que surge es si algunas empresas sabrán esperar hasta que emerja la segunda generación de científicos de datos y los aspirantes sean más numerosos, menos costosos y más fáciles de probar e incorporar en un entorno empresarial. ¿Por qué no dejarles el problema de buscar y domesticar profesionales exóticos a las startups y empresas de big data como GE y Walmart, cuyas estrategias agresivas exigen que estén a la vanguardia?

El problema de ese razonamiento es que el avance de los big data parece no detenerse. Si las empresas no aprovechan los inicios de esta tendencia por la falta de profesionales, se arriesgan a perder terreno como competidores y a que los socios de canal obtengan ventajas casi inalcanzables. Piense en los big data como una ola gigante que se está armando y que comienza a formar una enorme cresta. Si quiere atrapar esa ola, debe contar con personas que sepan surfear.

Reimpresión de HBR R1210D

En este momento, los científicos de datos son similares a los analistas cuantitativos de Wall Street de las décadas de 1980 y 1990

colegas de administración general tiene una posible desventaja. Interactuarán menos con especialistas similares, lo cual es necesario para que mantengan sus habilidades intactas y sus herramientas actualizadas. Los científicos de datos tienen que contactarse con comunidades de práctica, ya sea dentro de grandes empresas o de manera externa. Cada vez surgen más conferencias y asociaciones informales nuevas que apuntan a alentar la colaboración y el uso compartido de tecnología, por lo que las empresas deben alentar a los científicos a participar en ellas sabiendo que "una marea alta levanta todos los barcos".

Asimismo, los científicos de datos tienden a estar más motivados cuando se espera más de ellos. Los desafíos de acceder a los big data y estructurarlos a veces dejan poco tiempo o energía para los análisis sofisticados que implican optimización o predicción. Aun así, si los ejecutivos afirman que los informes simples no son suficientes, los científicos de datos dedicarán más esfuerzos a realizar análisis avanzados. Los big data no deben ser iguales a la "matemática simple".

El trabajo más popular de la década

A Hal Varian, el economista jefe de Google, se lo reconoce por la frase "el trabajo más solicitado de los próximos 10 años será el de estadístico. Algunos creen que estoy bromeando, pero ¿quién habría pensado que el trabajo más popular de la década de 1990 iba a ser el de ingeniero en informática?"

Si "popular" significa tener una gran demanda por sus cualidades exclusivas, entonces los científicos de datos ya son merecedores de ese título. Son difíciles y costosos de

ingeniería financiera, los cuales forjaron una segunda generación de profesionales que era más accesible para las empresas convencionales. Más tarde, en la década de 1990, se repitió el patrón con los ingenieros de búsqueda, cuyas habilidades intrincadas pronto comenzaron a enseñarse en los programas de informática.

La pregunta que surge es si algunas empresas sabrán esperar hasta que emerja la segunda generación de científicos de datos y los aspirantes sean más numerosos, menos costosos y más fáciles de probar e incorporar en un entorno empresarial. ¿Por qué no dejarles el problema de buscar y domesticar profesionales exóticos a las startups y empresas de big data como GE y Walmart, cuyas estrategias agresivas exigen que estén a la vanguardia?

El problema de ese razonamiento es que el avance de los big data parece no detenerse. Si las empresas no aprovechan los inicios de esta tendencia por la falta de profesionales, se arriesgan a perder terreno como competidores y a que los socios de canal obtengan ventajas casi inalcanzables. Piense en los big data como una ola gigante que se está armando y que comienza a formar una enorme cresta. Si quiere atrapar esa ola, debe contar con personas que sepan surfear.

Reimpresión de HBR R1210D

 Thomas H. Davenport es profesor invitado de la Escuela de negocios Harvard, asesor ejecutivo de Deloitte Analytics y coautor de *Judgment Calls* (Harvard Business Review Press, 2012). D.J. Patil, que anteriormente fue director de Productos de Datos de LinkedIn, es el científico de datos residente de Greylock Partners y el autor de *Data Jujitsu: The Art of Turning Data into Product* (O'Reilly Media, 2012).

Aviso de restricciones de uso de Harvard Business Review, mayo de 2009

El contenido de Harvard Business Review y de los boletines informativos de Publicaciones de la Escuela de negocios Harvard que se encuentra en EBSCOhost cuenta con una licencia para el uso particular privado de los usuarios autorizados de EBSCOhost. No está diseñado para el uso como material asignado de curso en instituciones académicas ni como material de capacitación o aprendizaje corporativo en empresas. Los titulares de licencias que pertenecen al ámbito académico no pueden usar este contenido en reservas electrónicas, paquetes de materiales electrónicos de lectura, vínculos permanentes en planes de estudios o por ningún otro medio que permita incorporar el contenido en recursos didácticos. Los titulares de licencias que pertenecen al ámbito empresarial no pueden alojar este contenido en sistemas de administración de aprendizaje ni utilizar vínculos permanentes o algún otro medio que permita incorporar el contenido en sistemas de administración de aprendizaje. El departamento de Publicaciones de la Escuela de negocios Harvard le concederá el permiso para que pueda establecer que este contenido esté disponible por dichos medios. Si desea hacer alguna consulta sobre el permiso y las tarifas, envíe un correo electrónico a permissions@harvardbusiness.org.