



CONVOCATORIA PARA ENTIDADES PÚBLICAS Y EMPRESAS PRIVADAS, PARA CONFORMAR EL BANCO DE RETOS APLICANDO CIENCIA DE DATOS. 2021

1. PRESENTACIÓN

Las empresas disponen un tesoro de datos sin aprovechar. La tecnología moderna ha permitido la creación y el almacenamiento de cantidades cada vez mayores de información. Se estima que el 90% de los datos en el mundo se crearon en los últimos dos años. Por ejemplo, los usuarios de Facebook suben 10 millones de fotos por hora. Pero estos datos frecuentemente solo permanecen almacenados en bases de datos y lagos de datos, básicamente sin tocar.

La gran cantidad de datos recopilados y almacenados por estas tecnologías puede generar beneficios transformadores para las organizaciones y sociedades de todo el mundo, pero solo si se saben interpretar.

La ciencia de datos revela tendencias y genera información que las empresas pueden utilizar para tomar mejores decisiones y crear productos y servicios más innovadores. Quizás lo más importante es que permite que los modelos de aprendizaje automático (ML) aprendan de las grandes cantidades de datos que se les suministran en vez de depender principalmente de los analistas de negocios para ver qué pueden descubrir a partir de los datos.

Los datos son la base de la innovación, pero su valor proviene de la información que los científicos pueden extraer y luego utilizar a partir de estos.

Muchas empresas han hecho de la ciencia de datos una prioridad y están realizando grandes inversiones en ella. En la última encuesta de Gartner ¹ a más de 3,000 CIO, los encuestados clasificaron la analítica y la inteligencia empresarial como las tecnologías de diferenciación más importantes para sus organizaciones. Los directores de informática encuestados ven estas tecnologías como las más estratégicas para sus empresas y están realizando las inversiones correspondientes.

Cada vez más las empresas se apoyan en las estadísticas, comportamientos e información recolectada de los clientes o el mercado para esbozar sus estrategias de negocio. Por esta razón, la analítica y los datos se han convertido en dos piezas fundamentales para las compañías, sin embargo, existe una escasez de talento humano en las empresas para exprimir el máximo rendimiento de los datos e impulsar el negocio.

De esta manera, este año MINTIC desarrollará un proceso de formación en ciencia de datos e inteligencia artificial denominado Data Science For All/Colombia (“DS4A”), el cual busca formar a 620 colombianos que deseen fortalecer sus habilidades digitales. Esta formación será desarrollada por Correlation One, con el apoyo de tres Instituciones de Educación Superior con Acreditación Institucional de Alta Calidad. Correlation One es líder mundial en el diseño e implementación de procesos de formación empresarial en Ciencia de Datos orientado a la fuerza laboral de alto nivel. Cuenta con más de 250,000 científicos de datos y más de 600 alianzas con

¹ <https://www.gartner.com/en/newsroom/press-releases/2018-01-25-gartner-says-self-service-analytics-and-bi-users-will-produce-more-analysis-than-data-scientists-will-by-2019>



los principales empleadores y universidades del mundo en Estados Unidos, China, Reino Unido, Canadá, Brasil, México, Argentina, Colombia e Irlanda. Los programas de formación desarrollados por Correlation One han sido incluidos en varios rankings y con estos han ayudado a fomentar el ecosistema de talento de IA en múltiples regiones de todo el mundo, han ganado reconocimiento mundial, incluyendo un Estudio de caso de Harvard Business School.

Correlation One ha diseñado programas sofisticados para empresas como: SoftBank, MinTIC, Citadel, JP Morgan, Point72, Two Sigma, MarketAxess, Twitch, ZS, Lyft, Deloitte, y la National Science Foundation y cuenta con un equipo de trabajo liderado por un profesor titular de Estadística en la Universidad de Harvard.

El modelo de formación utilizado por Correlation One es 40% teórico y 60% práctico. La parte práctica se desarrolla sobre casos reales o retos del sector público y privado, a través de los cuales se aplican los conocimientos y se brindan alternativas o soluciones a los requerimientos que surgen de sus procesos de transformación digital o cambios en su oferta de valor.

Durante el programa de formación, los participantes serán organizados en equipos, y cada equipo trabajará en el desarrollo de un reto específico. El reto es un desafío de ciencia de datos del mundo real de un sector particular. Con la dirección del equipo de enseñanza de nivel mundial, los participantes aplicarán sus aprendizajes durante el curso para desarrollar y dar solución a los retos seleccionados.

Por esta razón, MINTIC está interesado en conformar un banco de retos de entidades tanto públicas como privadas para ser desarrollados por los beneficiarios de este proceso de formación, y que los desarrollos respondan a las necesidades de algunas entidades tanto del orden nacional como regional que requieran soluciones a temas específicos de aplicación de Ciencia de Datos e Inteligencia Artificial.

2. OBJETIVO DE LA CONVOCATORIA

Seleccionar hasta 100 retos del sector público y privado para ser desarrollados aplicando ciencia de datos e inteligencia artificial, con el objetivo de brindar soluciones efectivas y pertinentes para la entidad, y así mismo, aportar a la formación integral de los estudiantes de la convocatoria de ciencia de datos DS4A.

3. CARACTERÍSTICAS DE LOS RETOS

- a) Deben ser problemas reales e identificados en cada entidad pública o empresa privada aplicante, que involucren el análisis de conjuntos de datos y que estén claramente definidos dentro del alcance de la ciencia de datos. Para conocer cómo identificar los problemas o retos que pueden ser resueltos aplicando ciencia de datos remítase a la infografía anexa "IDENTIFICACIÓN DE RETOS" y a los videos publicados en la convocatoria, donde Correlation One, entidad encargada de la formación en ciencia de datos les brinda una explicación de cómo identificar los retos que pueden ser resueltos a través de la ciencia de datos.



- b) No se admitirán propuestas de desarrollo de Software. Los resultados serán un análisis exploratorio, identificación de variables, comportamientos y patrones que permitan el diseño y visualización de la información para toma de decisiones.
- c) Los retos o casos deben contar con datos suficientes que sean relevantes para dar solución al problema planteado, para lo cual se firmará un acuerdo de confidencialidad (numeral 10), así mismo, deben estar disponibles para su consulta antes del 28 de mayo de 2021 y hasta el 3 de septiembre de 2021.

4. REQUISITOS PARA PARTICIPAR

- Ser una organización pública o privada nacional o territorial, entidad descentralizada o adscrita a una entidad pública.
- Adjuntar los documentos de representación legal y documentos de identidad de quien ejerza esta condición para la entidad o empresa. Para el caso de entidades públicas bastará que el Directivo que postula el reto remita su documento de identidad en los dos campos del formulario donde se solicita la cédula y el documento de representación legal.
- Presentar la autorización del representante legal por escrito para postular el reto, ver Anexo 2 Solicitud de inscripción. Para el caso de entidades públicas bastará que el Directivo que postula el reto suscriba la solicitud.
- Presentar el documento Anexo 1 Información del reto, donde debe detallar el reto, el resultado esperado y los datos con los que cuenta para que los estudiantes desarrollen el reto durante la formación.
- Designar a una o dos personas que pueden ser, funcionario y/o contratista y/o empleado de la entidad pública o empresa privada participante, uno para el rol de líder del reto y otro para el rol técnico, para que participen en las sesiones de trabajo programadas con el equipo técnico de la convocatoria. Es posible que una sola persona pueda ser líder y técnico del reto, si cuenta con los conocimientos necesarios para apoyar a los estudiantes en el desarrollo de éste.
- Sí la entidad pública o empresa privada presenta varios retos, cada uno debe contar con un líder de reto y un técnico designado.
- Disponer y remitir la información necesaria cumpliendo con las características requeridas del equipo técnico en el marco de la convocatoria.
- El líder o técnico designado por la entidad o empresa debe brindar la información de manera oportuna a los equipos de estudiantes.

5. CRONOGRAMA DE LA CONVOCATORIA

ACTIVIDADES	FECHA INICIO	FECHA FIN
Inscripción	16 de abril	7 de mayo (4:00 p.m.)
Envío de conjunto de datos a Correlation One	8 de mayo	21 de mayo
Revisión y selección de los retos	8 de mayo	27 de mayo
Publicación de resultados	28 de mayo	28 de mayo



6. PROCESO DE POSTULACIÓN

- Deben identificar los retos que consideren pueden ser resueltos a través de la aplicación de la ciencia de datos y que cuenten con la información necesaria para que los participantes de la convocatoria de ciencia de datos puedan desarrollarlos y brindarles una posible solución. Para llevar a cabo esta identificación de los retos deben revisar la infografía y los videos anexos, donde se dan pautas claras para lograr identificar qué tipos de retos pueden ser desarrollados aplicando ciencia de datos.
- Enviar la cantidad de retos que deseen postular, para lo cual deben diligenciar un formulario para cada reto dentro de los plazos establecidos. La inscripción a partir del 12 de abril de 2021 debe realizarse a través del link: <https://forms.mintic.gov.co/cientificosdatos/retos>
- Diligenciar el Anexo 1 Información del reto y Anexo 2 Solicitud de inscripción, donde se debe describir detalladamente el problema a resolver y se debe brindar la autorización del representante legal de la empresa o el directivo a cargo para el caso de las entidades públicas, para lo cual puede revisar los videos anexos.
- Los retos enviados después de la fecha y hora indicada en el calendario de la convocatoria y por medio distinto al establecido, no se tendrán en cuenta dentro del proceso de selección del banco de retos. Una vez inscrito no se podrán hacer modificaciones a los retos y formularios.
- En dado caso que al momento de inscribirse no remita todos los datos solicitados, Correlation One se comunicará con ustedes para solicitarles remitir toda la información y para apoyarlos en la identificación de estos datos.
- Una vez cierre la etapa de inscripción de retos el 3 de mayo, Correlation One remitirá un correo electrónico solicitando el conjunto de datos que usted indicó en su postulación y le dará las pautas para remitir esta información. Para enviar este conjunto de datos tendrán plazo hasta el 16 de mayo.
- El equipo de científicos de datos de Correlation One luego revisará los conjuntos de datos enviados para confirmar que sean suficientes y apropiados para el proyecto propuesto. De ser necesario Correlation One se comunicará con ustedes para solicitarles información adicional a la remitida.

7. SELECCIÓN DE RETOS

Serán seleccionados los retos que cumplan con los requisitos establecidos para formar parte del banco de retos aplicando ciencia de datos. En ese sentido, Correlation One entidad encargada de la formación de ciencia de datos, será quien realizará la selección de los retos de acuerdo con la información remitida por la empresa, sin embargo, si identifican que el problema o reto postulado se puede resolver a través de la ciencia de datos, pero no cuenta con la información o los datos requeridos para ser desarrollado, se comunicarán con la entidad para solicitar envío de esta información requerida. Por lo tanto, es muy importante que la persona líder



representante del reto esté pendiente de su correo durante el proceso de revisión y selección de los retos.

La evaluación realizada por Correlation One tendrá en cuenta los siguientes elementos:

- El reto debe cumplir con las características técnicas descritas en el Numeral 3 del presente documento.
 - a) Deben ser problemas reales que involucren el análisis de conjuntos de datos y que estén claramente definidos dentro del alcance de la ciencia de datos.
 - b) No se admitirán propuestas de desarrollo de Software
 - c) Los retos o casos deben contar con datos suficientes que sean relevantes para dar solución al problema planteado.
- Diligenciamiento del formulario en su totalidad.
- Diligenciamiento de los dos anexos.
- Envío de los documentos del representante legal o del directivo a cargo del reto para el caso de las entidades públicas.
- Disponibilidad de los datos e información relacionada con el reto.

Los retos seleccionados, serán notificados a las personas de contacto relacionadas en el formulario, con quienes se establecerá comunicación directa a fin de suscribir las alianzas necesarias y definir las políticas de derechos de autor que se apliquen a este caso particular.

Los retos que no cuenten con la información y datos necesarios para su desarrollo no serán seleccionados dentro del banco de retos.

8. DESARROLLO DE LOS RETOS

- Los retos seleccionados serán entregados a los participantes de la formación de ciencia de datos DS4A para que escojan cuál desarrollarán durante el curso, por lo tanto, un reto puede ser resuelto por más de un equipo. En este sentido, la entidad es quien decidirá con cuántos equipos tiene la capacidad de trabajar.
- Equipos de 5 a 7 participantes en el programa de formación de ciencia de datos DS4A trabajarán en un reto durante 12 semanas de clase. Aplicarán ciencia de datos y las técnicas de inteligencia artificial que aprendan durante el curso, con el asesoramiento de un equipo de instructores de clase mundial, para crear una solución impactante que pueda ser implementada por la entidad.
- Cada equipo contará con una persona de contacto, que será quien se comunicará con las entidades o empresas para solicitar información y llevar a cabo el desarrollo del reto.
- Las entidades que sus retos sean seleccionados dentro del banco de retos, deberán firmar un acuerdo de confidencialidad con Correlation One, así mismo, cada equipo de DS4A que vaya a desarrollar su reto firmará un acuerdo de confidencialidad con Correlation One.
- Posteriormente, las entidades deberán remitir a los equipos que los seleccionaron toda la información y documentos necesarios para dar una solución al reto propuesto.



- Los equipos estarán en permanente contacto con las entidades y empresas a través de todo el proceso de desarrollo del reto, por lo que, el líder y técnico del reto debe estar dispuesto a poder reunirse con los equipos de estudiantes.

9. LÍDER Y TÉCNICO DEL PROYECTO

Es importante designar un punto de contacto en su empresa o entidad que pueda actuar como Líder o guía para los equipos del proyecto. Los equipos tendrán preguntas sobre los conjuntos de datos o sobre su negocio, y un mentor receptivo se asegurará de que aproveche al máximo sus equipos. También es una excelente manera de persuadir a los candidatos sobre los méritos de su organización, en caso de que desee contratar graduados de DS4A.

El técnico será la persona que más conozca a profundidad los datos e información que se necesiten para dar solución al reto, ya que de esto depende que el desarrollo del reto por parte de los estudiantes sea el esperado por la empresa y pueda llegarse a implementar por parte de ésta.

10. CONFIDENCIALIDAD

Todos los participantes de la formación de ciencia de datos DS4A firmarán acuerdos de confidencialidad y no divulgación con Correlation One para proteger la privacidad de los datos. No se les permitirá usar los conjuntos de datos fuera del programa de capacitación. Si sus datos tienen información personal confidencial, es necesario anonimizar los campos para la protección de la privacidad y aplicación de la ley de protección de datos.

Así mismo, Correlation One firmará un acuerdo de confidencialidad con cada una de las entidades públicas o empresas privadas con el objetivo de proteger la información suministrada por las mismas.

De requerirse, los estudiantes que desarrollen su reto pueden firmar directamente un acuerdo de confidencialidad con la entidad.

11. RESULTADOS

- Cada equipo creará una aplicación que resuelva el problema de ciencia de datos de la entidad, dependiendo de los datos suministrados y el acompañamiento que reciban del líder y técnico del reto. Como mínimo, la aplicación tendrá los siguientes tres componentes técnicos: interfaz de usuario front-end, canalización de datos, y herramientas & modelos analíticos. Los componentes específicos de la aplicación final dependerán en gran medida de los detalles específicos del proyecto.
- Informe de 10-20 páginas que detalla los objetivos, métodos de análisis y conclusiones del proyecto.
- Presentación en vídeo de 10 minutos que muestra los aspectos más destacados del proyecto.
- Datafolio que resume de forma visual y concisa los resultados en una página.



- Una vez se haga entrega de los resultados, la empresa deberá firmar un acta de recibido que será remitida por Correlation One.

12. OBSERVACIONES ESPECIALES

MinTIC y Correlation One estarán informando de los cambios que sucedan en el desarrollo de la presente Convocatoria y que se deriven de eventos fortuitos que no corresponden a la voluntad de estas entidades, por lo cual las entidades no se hacen responsables.

Es importante resaltar que al finalizar la formación en ciencia de datos DS4A se llevará a cabo un evento para realizar networking con diferentes empresas del sector público y privado, con el objetivo de que puedan contratar a los graduados de la formación de ciencia de datos DS4A, en ese sentido, los invitamos a participar.

Si presenta alguna inquietud sobre el proceso de selección o de formación, la persona interesada en participar puede comunicarse al correo electrónico: info.talentodigital@mintic.gov.co.