**Formulario Evaluación Data Reto Integrativo**

Muchas gracias por participar en el proceso de evaluación de organizaciones para el Reto Integrativo.

Por favor les pedimos hacernos llegar los siguientes documentos al siguiente correo: [mburgos@aporta.org.pe](mailto:mburgos@aporta.org.pe)

* Formulario Evaluación de Data Reto Integrativo completo
* 1 muestra de cada fuente de datos presentada en el formulario (debe mostrar los campos principales y 5 registros mínimo, anonimizada, puede ser en CSV o Excel)
* Diccionario de datos, de no contar con uno les pedimos una breve descripción de variables que puedan ser difíciles de entender al revisar la tabla (p.ej. perfiles, categorías internas, otros)
* Brochures, links a información relevante sobre los proyectos activos de su organización

Si tienen alguna consulta, no duden en comunicarse con nosotros.

# Datos de contacto

|  |  |
| --- | --- |
| **Nombre de la organización** | Perú Champs |
| **Principales proyectos/servicios activos** | Brindar una educación de calidad a niños talentosos con bajos recursos, y que no tienen la oportunidad de acceder a escuelas que desarrollen su potencial. Los estudiantes becados -al 75% por Perú Champs- tienen la oportunidad de estudiar en Innova Schools, reconocida como la red de escuelas más grande de nuestro país, con acreditación internacional AdvancED que certifica el más elevado estándar de calidad. Además, la beca ofrece oportunidades adicionales para el desarrollo de nuestros beneficiarios, incluyendo educación complementaria en formación de liderazgo y apoyo psicológico continuo. |
| **Nombre persona de contacto** |  |
| **Correo contacto** |  |
| **Celular contacto** |  |

# Descripción de principales fuentes de datos

* 1. Describa brevemente las principales fuentes de datos de la organización/proyecto/servicio (p.ej: Datos demográficos de beneficiarios)

Fuente de datos 1

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre fuente de datos | BD Champs | | |
| Breve descripción | Base de datos con información personal de nuestros beneficiarios | | |
| Fecha creación | 2014 | Última fecha actualización | 07/02/2024 |
| Frecuencia de actualización | Diaria | Esta frecuencia es regular | Sí |
| Breve descripción del procedimiento del recojo de la data (p.ej. si es por encuesta, dar información de  ficha técnica, si es transaccional explicar) | Parte de la información proviene de un formulario de postulación (datos personales y de contacto) y otra parte de la información proviene directamente de Innova (status de matrícula, sede, pagos), solo una columna viene de AWS que verifica si se completó el convenio en el Intranet. | | |
| Número aprox. total de registros | 2258 registros | | |
| Tipo de datos (p.ej. fotos, videos, texto) | Texto | | |

Fuente de datos 2

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre fuente de datos | BD Renovación | | |
| Breve descripción | Es una base de datos que compila la información relevante para renovar o no la beca, por ejemplo: Información de conducta, nota, NSE, Blacklist, datos de donantes directos y morosidad. | | |
| Fecha creación | 2015 | Última fecha actualización | 04/01/2024 |
| Frecuencia de actualización | Anual | Esta frecuencia es regular | No |
| Breve descripción del procedimiento del recojo de la data (p.ej. si es por encuesta, dar información de  ficha técnica, si es transaccional explicar) | La principal fuente de información que alimenta la BD es la proporcionada por Innova Schools (Notas, Conducta, Morosidad), equipo de Trabajo Social (NSE), equipo de Selección y Seguimiento (Blacklist), equipo de Fundraising (Donantes) | | |
| Número aprox. total de registros | 1796 | | |
| Tipo de datos (p.ej. fotos, videos, texto) | Texto | | |

1. Describa brevemente las principales fuentes de datos de la organización/proyecto/servicio (p.ej: Datos demográficos de beneficiarios)

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre fuente de datos | Liderazgo y Consultoría (Programa de formación en liderazgo) | | |
| Breve descripción | Es la BD del área de Programa de Liderazgo cuyo fin es contar con información personal de los beneficiarios, información derivada del área de psicología, innova y de la misma área. El área desarrolla programas de fortalecimiento de habilidades de liderazgo durante el año y mediante la bd se asegura de cumplir con su cobertura anual, a su vez, evitar los cruces entre programas. | | |
| Fecha creación | 2022 | Última fecha actualización | 06/02/2024 |
| Frecuencia de actualización | Semanal | Esta frecuencia es regular | Sí |
| Breve descripción del procedimiento del recojo de la data (p.ej. si es por encuesta, dar información de  ficha técnica, si es transaccional explicar) | Se traslada información de la BD Champs (que contiene información general de los beneficiarios), información del área de psicología ( si cuentan con carta de compromiso para su renovación, el nivel de riesgo psicológico) e información de la misma área que se completa según la participación en los diversos programas. | | |
| Número aprox. total de registros | 2082 | | |
| Tipo de datos (p.ej. fotos, videos, texto) | texto | | |

Fuente de datos 3

Fuente de datos 4

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre fuente de datos | NRE (Nivel de Riesgo Psicológico) | | |
| Breve descripción | Es la BD que contiene toda la información psicológica derivada de encuestas y entrevistas entre los beneficiarios y sus padres con el equipo de psicología de la organización. | | |
| Fecha creación | 2022 | Última fecha actualización | 02/02/2024 |
| Frecuencia de actualización | Semanal | Esta frecuencia es regular | Sí |
| Breve descripción del procedimiento del recojo de la data (p.ej. si es por encuesta, dar información de  ficha técnica, si es transaccional explicar) | Toda la información derivada de encuestas y entrevistas se conglomera en la presente BD, es una fuente de información valiosa, ya que, en base a los niveles de riesgos que se establecen para nuestros beneficiarios, se diferencian las acciones hacía ellos como por ejemplo: frecuencia de intervenciones psicológicas, asignación de donante, participación en programas de liderazgo, etc. | | |
| Número aprox. total de registros | 1902 | | |
| Tipo de datos (p.ej. fotos, videos, texto) | texto | | |

1. Describa brevemente las principales fuentes de datos de la organización/proyecto/servicio (p.ej: Datos demográficos de beneficiarios)

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre fuente de datos | BD NSE (Nivel Socioeconómico) | | |
| Breve descripción | Es la BD que contiene la información Socioeconomica de nuestros beneficiarios y que nos permite tomar decisiones sobre la renovación de la beca. | | |
| Fecha creación | 2016 | Última fecha actualización | 25/01/2024 |
| Frecuencia de actualización | Mensual | Esta frecuencia es regular | No |
| Breve descripción del procedimiento del recojo de la data (p.ej. si es por encuesta, dar información de  ficha técnica, si es transaccional explicar) | Sobre la recolección de información, se realiza principalmente en base a la información obtenida en entrevistas por parte del área del área de Trabajo Social, encuestas enviadas por nosotros e información derivada por parte de Financiera Oh! | | |
| Número aprox. total de registros | 1797 | | |
| Tipo de datos (p.ej. fotos, videos, texto) | Texto | | |

Fuente de datos 5

Fuente de datos 6

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre fuente de datos | Ingresar | | |
| Breve descripción | Ingresar | | |
| Fecha creación | Desde qué año existe esta fuente de datos | Última fecha actualización | Año última actualización |
| Frecuencia de actualización | Ingresar | Esta frecuencia es regular | Sí/No (explicar) |
| Breve descripción del procedimiento del recojo de la data (p.ej. si es por encuesta, dar información de  ficha técnica, si es transaccional explicar) | Ingresar | | |
| Número aprox. total de registros | Ingresar | | |
| Tipo de datos (p.ej. fotos, videos, texto) | Ingresar | | |

**\*** Si cuentan con más fuentes de datos, pueden llenar otro formulario incluyendo las demás fuentes.

* 1. ¿Las fuentes de datos se encuentran relacionadas? ¿Cómo?

Algunas fuentes de datos se encuentran relacionadas mediante fórmula IMPORTRANGE.

* 1. ¿Además de estas fuentes de datos, tienen otro tipo de información no estructurada almacenada digitalmente que sea relevante para sus proyectos/servicios? (p.ej. transcripciones, grabaciones, e-mails, fotografías)?

Contamos con fotografías de nuestros beneficiarios pero que no se encuentran vinculadas a las bd’s.

* 1. ¿Existe alguna fuente de datos externa que sea relevante para su proyecto que se pueda disponer para el proyecto? (p.ej. data proporcionada por el MINEDU o alguna organización externa a ustedes)

Las proporcionadas por Innova Schools y Financiera Oh!

* 1. ¿Hay algún proyecto en los próximos 6 meses que vaya a cambiar de forma importante la estructura, la forma de recojo o características de los datos? Por favor explique el proyecto y el impacto que tendría en los datos

No.

* 1. ¿Alguna especificación adicional que considere relevante sobre las fuentes declaradas?

No.

# Sobre las posibles soluciones a desarrollar en el Reto Integrativo

De ingresar al Reto Integrativo, el equipo de nuestros alumnos, profesionales relacionados a analítica de datos diseñaran en conjunto con los representantes de su organización la solución a desarrollar. Sin embargo, de manera preliminar quisiéramos saber:

¿Qué necesidades/dolores actuales de su organización cree que podría ayudar nuestro equipo de alumnos y expertos a aliviar durante el Reto Integrativo? ¿Tienen algunas soluciones potenciales en mente que creen se podrían trabajar (por favor listarlos)? (esta es una pregunta de levantamiento preliminar, se pueden plantear ideas o necesidades generales, no es requisito tener una solución definida para ingresar al Reto)

1. Al realizar una revisión total de las BD’s con motivo del formulario de postulación para el Reto Integrativo hemos podido apreciar que uno de los principales dolores de la organización es contar con una Base de Datos general más detallada que pueda ser nutrida –relacionada- a las bd’s que gestiona cada uno de los equipos. Una vez relacionada y centralizada toda la información, resultaría imprescindible contar con dos soluciones conjuntas: una plataforma de consultas en donde se pueda revisar toda la información de un beneficiario –o un grupo de beneficiarios- y un dashboard que nos permita analizar la información a nivel mínimo descriptivo sobre el avance de las áreas que tienen contacto con nuestros beneficiarios y que, también, cuenta con información y distribuciones básicas del perfil de nuestros beneficiarios.
2. Realizar un estudio de línea base –con información pública disponible- sobre la diferencia académica existente sobre estudiantes que son beneficiarios de la organización y de quienes no lo son.
3. Desarrollar métricas y establecer un sistema para monitorear y evaluar el impacto de las becas en la vida de nuestros beneficiarios. Esto podría incluir indicadores como el rendimiento académico, la conducta escolar, la participación en actividades extracurriculares, etc.
4. Utilizar técnicas de análisis de datos para segmentar a los beneficiarios en grupos con características similares. Esto podría ayudar a personalizar los programas y servicios ofrecidos a cada grupo, maximizando así el impacto de la educación complementaria que ofrecemos.
5. Utilizar técnicas de análisis de datos para identificar tendencias a lo largo del tiempo y predecir posibles

cambios en las necesidades de nuestros beneficiarios. Esto podría ayudar a adaptar los programas y servicios de manera proactiva.

1. Evaluación de la organización

Marque el nivel de desarrollo en el que se encuentra su organización/proyecto en cada dimensión

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimensión** | **Nivel** | | | |
| **Básico** | **Intermedio** | **Avanzado** | **Líder** |
| **Cultura de datos** | ☐  El staff sabe qué datos hay pero no reconocen su importancia. | X  Pocos individuos saben qué data hay y qué se puede hacer con ella. | ☐  La organización tiene una idea clara de cómo los datos pueden servir para tomar decisiones de negocio | ☐  La organización tiene una cultura de datos y demanda que las decisiones se justifiquen con datos |
|  | ☐ | X | ☐ | ☐ |
| **Recolección de datos** | El staff provee datos esporádicamente porque están obligados a hacerlo. | El staff regularmente provee datos porque están obligados a hacerlo. | El staff regularmente provee datos y genera insights | El staff provee datos en tiempo real y decide en función a los insights |
|  | ☐ | X | ☐ | ☐ |
| **Importancia para el liderazgo** | Los líderes no entienden cómo los datos pueden ayudar a cumplir la misión de la organización. | Los líderes quieren utilizar los datos pero no saben cómo exactamente | Los líderes tienen una idea clara de cómo los datos pueden dirigir las decisiones de negocio. | Los líderes crean una cultura de datos en la organización y demandan que las decisiones se justifiquen con datos |
| **Recursos humanos** | ☐  El staff mantiene sus datos en silos. | X  El staff puede ver el valor de los datos pero no tiene la capacidad para crear insights valiosos. | ☐  El staff ve bien el valor de los datos, pero no tiene las habilidades suficientes o infrastructura para  crear insights valiosos | ☐  La organización tiene un staff dedicado a la gestión de la data, almacenamiento, rigor y limpieza. |
|  | ☐ | X | ☐ | ☐ |
| **Políticas de uso de datos** | No existen políticas para el uso, transferencia y compartir de la data | Hay políticas para el uso, transferencia y compartir de la data pero no cubre toda la data que existe | Hay buenas políticas para el uso, transferencia y compartir de la data internamente. | Hay buenas políticas para el uso, transferencia y compartir de la data interna y externamente |
|  | ☐ | ☐ | X | ☐ |
| **Probabilidad de aplicación de soluciones usando data analytics** | Es poco probable que comencemos a usar data analytics a partir de esta experiencia | Es probable que incorporemos análisis básicos de data analytics a partir de esta experiencia | Es probable que esta experiencia nos ayude a alcanzar formas más avanzadas de gestionar la data | Esta experiencia aportará una solución más de las varias soluciones que tenemos utilizando data science |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dimensión** | **Nivel** | | | |
| **Básico** | **Intermedio** | **Avanzado** | **Líder** |
| **Accesibilidad** | ☐  La data solo es accesible desde la fuente que la genera | ☐  La data es accesible desde fuera, pero en un formato que nos genera algunos problemas | X  La data es accesible desde fuera con un formato estándar | ☐  La data es accesible desde un API en un formato estándar |
| **Almacenamiento** | ☐  Papel | ☐  PDF, Excel o bases de datos con formato propietario | X  Bases de datos y CSV, JSON, TXT o archivos de texto plano | ☐  Bases de datos que permiten llamadas desde un API |
| **Integración** | ☐  La data sólo se encuentra donde se genera | X  La data se exporta ocasionalmente | ☐  La data se agrega en un data warehouse | ☐  La data se integra con información externa a un data warehouse |
| **Relevancia y suficiencia** | ☐  La data es irrelevante al objetivo de proyecto | ☐  La data es relevante, pero es inconsistente | X  La data es relevante pero no es suficiente | ☐  La data es relevante y suficiente |
| **Calidad** | ☐  La data tiene muchos valores vacíos, filas vacías y errores | ☐  La data tiene filas vacías y posibles errores | ☐  La data tiene pocos valores vacíos y pocos errores | X  La data no tiene valores vacíos, baja posibilidad de error y procesos de validación |
| **Datos temporales** | ☐  Solo se mantiene la data actual, no se acumula data histórica | ☐  La data histórica es almacenada, pero se sobre- escriben algunos valores | X  La data histórica es almacenada y la nueva se agrega a la antigua | ☐  La data histórica es almacenada y los schemas se mapean de manera consistente |
| **Documentación** | X  No hay documentación de las variables, pero se pueden crear diccionarios | ☐  Ya existen diccionarios | ☐  Los diccionarios tienen información sobre las condiciones bajo las que se capturó la data | ☐  Los diccionarios tienen información sobre los supuestos y posibles sesgos de la data |