

**Conectando oportunidades en la era digital**  
**Consultoría**  
**Estructura Tecnológica y Metodológica para**  
**Laboratorio de Transformación Digital para MIPYMES**

Dudas y aclaraciones

Con el objetivo de recibir propuestas en línea con lo que se propone nuestro proyecto y aclarar dudas que van surgiendo en el camino sobre los Términos de Referencia publicados, les compartimos algunas preguntas que nos han llegado de los postulantes.

1. ¿Cuentan con información histórica sobre cursos realizados y resultados de éstos para identificar el impacto en los participantes?

No. No se han realizado mediciones de impacto sobre los cursos anteriormente.

2. ¿Cuentan con información histórica relevante (financiera, demográfica, madurez digital, descriptiva, etc.) para ser utilizada como input para los algoritmos predictivos?

No. La idea es que se comience a registrar esta información para ser analizada a futuro. En un principio, esta información podrá ser utilizada sólo para análisis estadístico, para luego poder ser utilizada para entregar algoritmos predictivos

3. ¿Cuentan con documentación sobre la metodología utilizada en las mediciones que ya realizan?

No, pero nuestro equipo está elaborando un documento de marco teórico que describe la documentación asociada a la metodología relacionada a los conceptos principales del proyecto

4. ¿Se mantiene el presupuesto de USD 60.0000 para el desarrollo del proyecto?

No, el presupuesto se deja a libre propuesta teniendo en cuenta el valor de mercado de una consultoría como esta.

5. ¿Cuáles son los criterios de empleabilidad y productividad que utilizan o esperan para realizar los análisis?

Definir estos criterios es parte de la metodología. Aún así, basados en la definición de la Organización Internacional del Trabajo (OIT), entendemos por empleabilidad la aptitud de la persona para encontrar y conservar un trabajo, para progresar en el trabajo y para adaptarse al cambio a lo largo de su vida profesional. Por tanto, es un concepto en el que se relacionan el aprendizaje, el trabajo y las empresas.

Por otra parte, entendemos el concepto de productividad de acuerdo a la forma en que lo utiliza el Ministerio de Economía en la realización de su Encuesta Longitudinal de Empresas (ELE), donde la define como una medida de la eficiencia en la producción, generalmente expresada como una relación entre los bienes producidos y los recursos utilizados para este fin.

6. En el documento se mencionan fuentes externas. ¿Éstas fuentes externas son información provista por entes como por ejemplo el INE, Centro de estudios de la Cámara de Comercio, Gremios, etc.)? ¿Nos las pueden listar?

Efectivamente son fuentes de instituciones públicas y privadas con las que hemos tenido conversaciones y estamos en proceso de alianzas.

CORFO

Ministerio de Economía

Instituto Nacional de Estadísticas

Servicio de Impuestos Internos

Cámara Chilena de la Construcción

Cámara de Comercio

Ministerio del Trabajo

Ministerio de Agricultura

SENCE

Sernatur

Ministerio de Transporte y Telecomunicaciones

### **Preguntas Respecto al Objetivo del Proyecto:**

1. En base al conocimiento del rubro o sector en el cual se desarrollará el proyecto:

- ¿Cuáles son las fuentes de información internas y/o externas válidas para la obtención de los datos que permitan las definiciones de?:

Las fuentes serán captura propia a través de herramientas de diagnóstico, evaluaciones y utilización de información de organismo públicos que tienen relación con los análisis que queremos realizar en el proyecto, tales como, SII, INE, Minecon, CORFO, Servicio de Desarrollo Social, Aduanas, etc.

- Las ocupaciones y habilidades demandadas en el contexto MIPYMES: elaboración propia
- Oferta de formación, impartidas por proveedores de formación validados por el solicitante y/o entidad reguladora correspondiente: contratamos líneas de formación en el marco del proyecto para evaluar la herramienta y realizar las primeras iteraciones de las mediciones

2. ¿Cuáles son los contenidos mínimos a considerar respecto a la identificación y desarrollo de herramientas y cursos de formación para MIPYMES?

Toda esta información se está preparando para entregar el marco teórico de investigación que tendrá el proyecto de acuerdo a contenidos, metodologías, perfiles de los relatores y alumnos, etc.

- ¿Qué estudios y análisis complementarios deben robustecer estos contenidos? todo será en base a iteraciones por lo que esperamos que el marco teórico sea la base para ir mejorando la herramienta de análisis a través de la investigación en base a datos

3. ¿Cuáles son los indicadores y/o KPIs claves para cada sector (público, privado y de proveedores de formación) para la toma de decisiones referidas a la oferta de formación misma?

Los KPI actualmente están entregados en relación a números de Mipymes formadas, el análisis de aprendizaje es lo que vamos a revisar nosotros como una herramienta que pueda servir a las líneas de formación a mejorar contenidos u otros aspectos que permitan terminar con un aprendizaje efectivo y aplicado para que las Mipymes.

Nosotros queremos medir aprendizaje efectivo, lo cual se deberá demostrar a través de la aplicación de una herramienta de diagnóstico de entrada y de salida, más herramientas de evaluación y otras herramientas de medición que estamos construyendo, para hacer mediciones de corto y mediano plazo.

4. ¿Qué herramientas de desarrollo de capital humano, habilidades digitales, adaptabilidad a nuevos mercados y otros factores necesarios a considerar deben ser considerados como mínimo para medir los índices de competitividad y productividad de las MIPYMES?

- ¿Cómo se mide la competitividad y productividad en las MIPYMES?

Entendemos el concepto de productividad de acuerdo a la forma en que lo utiliza el Ministerio de Economía en la realización de su Encuesta Longitudinal de Empresas (ELE), donde la define como una medida de la eficiencia en la producción, generalmente expresada como una relación entre los bienes producidos y los recursos utilizados para este fin.

- ¿Cuáles son los rangos de medición para competitividad y productividad de las MIPYMES? Se mide de acuerdo a ventas y número de trabajadores en un período de tiempo.

### **Preguntas Respecto a las Especificaciones de la Consultoría, Servicios Requeridos:**

1. Para la elaboración del benchmark con otras experiencias internacionales en la medición del impacto de cursos e instancias de formación tecnológica en la transformación digital y empleabilidad de MIPYMES. Con el supuesto de tener las fuentes de información y KPIs claves para la medición nacional y que deberemos utilizar los mismos criterios para el benchmark.
  - ¿Qué fuentes de información se utilizarán para la homologación de estos indicadores a nivel internacional?

La idea es conocer otras plataformas que trabajan con investigación de las mipymes ojalá con el impacto de la Transformación tecnológica, Madurez Digital y/o impacto de formación para la adopción tecnológica, ya sean creaciones públicas o privadas, como por ejemplo el Centro de Transformación Digital Empresarial de Colombia.

- ¿Qué países se utilizarán como referencia?

Puede ser cualquier país que lo esté desarrollando,

2. Considerando la definición para una metodología de medición de impacto de cursos e instancias de formación:
  - ¿Cuáles son los atributos o indicadores a-priori, intermedios y a-posteriori a considerar para las MIPYMES en el diseño e implementación de dicha metodología?

Esto es parte de la propuesta metodológica, por lo que es abierto y depende de la metodología definida por los proponentes. Por nuestra parte, también propondremos indicadores para estas mediciones, los cuales serán indicados en el documento de marco teórico.

3. Del diseño e implementación del algoritmo predictivo, en base a datos históricos de medición de impacto, que permita recomendar a una MIPYME el curso o instancia de formación con mayor impacto esperado en base a su diagnóstico a-priori.

- ¿Qué score o ponderación deberán configurarse para definir que un curso o instancia de formación es bueno o malo? se medirá de acuerdo a aprendizajes adquiridos, lo estamos definiendo.
- ¿Deben ser variables de decisión dicotómicas, o se pueden extender un score o categorías para este punto?

Las variables de decisión no tienen restricción ya que dependen del tipo de algoritmo a utilizar: por ejemplo, algunas redes neuronales o SVM se beneficiarían de estimaciones de impacto numéricas mientras que ciertos árboles de decisión podrían funcionar mejor con variables categóricas o binarias/dicotómicas. Se debe hacer énfasis en la exactitud (accuracy) y precisión del algoritmo.

4. Del diseño e implementación del sistema de alerta temprana basado en un algoritmo predictivo que, en base a datos históricos de medición de impacto, permita detectar dentro de cada curso o instancia de formación, aquellas MIPYMEs con un impacto esperado bajo, en base a su diagnóstico a-priori y sus mediciones intermedias. Mismas observaciones de puntos anteriores:
  - ¿Qué score o ponderación deberán configurarse para definir un impacto bajo en este punto?

Para ambos algoritmos predictivos, la clase o *target* a predecir es el impacto de las líneas de formación en la MIPYME. Esta medición de impacto es uno de los elementos centrales de la propuesta y su medición debe ser elaborada por los proponentes. De todas maneras, nuestro equipo está elaborando de forma independiente un esquema de medición de impacto que será compartido con la empresa que adjudique el proyecto a través del documento de marco teórico.

La diferencia entre los algoritmos predictivos de recomendación y de alerta temprana es que el segundo debe considerar mediciones intermedias durante las líneas de formación en su estimación de impacto, y debe realizar estimaciones de impacto de forma periódica.

5. Del diseño e implementación del portal web para el Laboratorio (front-end), que permitirá y regulará el acceso a herramientas de captura de datos de diagnóstico, mediciones intermedias, acceso a datos, predicciones, publicaciones del laboratorio, gestión de contenidos, medios y RR.SS., etc;:
  - ¿Qué formatos de inputs como mínimo deben ser considerados?

Los inputs corresponden a todos los instrumentos y herramientas de medición en todas sus fases que sean necesarios para la medición de impacto, a determinar en su formato y forma según la metodología elegida y definida por cada proponente

- Se entiende que dependerá de las fuentes de orígenes de los datos, pero como mínimo debemos considerar API's, CSV, ¿otros?

Si, además deben considerar conexiones a bases de datos externas (sql y nosql), servicios web (SOAP, REST) otros tipos de archivos (xml, html, txt)

- ¿Qué RRSS como mínimo se deben considerar; LinkedIn, Twitter?, ¿otros?

Con eso es suficiente

6. Del diseño e implementación de una arquitectura tecnológica integrada (back-end) que permita la captura de datos desde diversas fuentes, su preprocesamiento y limpieza, almacenamiento, análisis estadístico y predictivo, visualización de información y de los resultados de análisis.
- ¿Existe algún requerimiento de tecnología específico, infraestructura en la nube, o dedicada o híbrida requerida?

No hay ningún requerimiento particular, la arquitectura puede alojarse en cualquiera de estos tres tipos de infraestructura

- ¿Algún requerimiento particular con la residencia de los datos?

Tampoco hay un requerimiento en particular en este punto. Se están elaborando políticas de gobernanza de datos, lo cual no está finalizado.

### **Preguntas Respecto a Resultados Esperados**

1. De las propuestas de valor orientadas a los diferentes stakeholders del proyecto (i.e., Laboratorio, MIPYMEs, BID, ASECH, Gobierno, ONGs, centros de investigación, etc) identificando sus necesidades, preguntas y respuestas requeridas para generar interés de la comunidad:
  - ¿Tenemos claro y definido todos y cada uno de los stakeholders involucrados en el alcance del proyecto con sus definiciones y responsabilidades asociadas? **De lo contrario no podremos cumplir con lo solicitado en este punto.**

Si tenemos definidos los stakeholders del proyecto a nivel Nacional que queremos llegar en primera instancia con el proyecto y perfiles a los que queremos llegar también que se vayan involucrando con posterioridad de acuerdo con los intereses y propósitos que estos tienen.

2. Respecto al portal web para el Laboratorio diseñado en base a los requerimientos de análisis de los diferentes stakeholders , implementado y aprobado, que permita la captura de todos los datos necesarios para la medición y predicción de impacto de cursos e instancias de formación, y que cumpla con las obligaciones de la Ley 19.628 sobre Protección de la vida privada.
  - ¿Existe algún requerimiento tecnológico específico, referido a framework u otro relacionado?
  - ¿Existe definiciones respecto al hosting de dicho portal? estamos trabajando en ello, lo tendremos antes de que se empiece a implementar la infra tec.
  - ¿Cuál es el alcance general y en detalle de la Ley 19.628 sobre Protección de la vida privada, tenemos dicha cobertura con nuestro actual portafolio de proyectos desarrollados? Estamos trabajando con abogados especialistas para el desarrollo de la gobernanza de datos para incorporar estándares internacionales al desarrollo para trabajar la privacidad por diseño y uso ético del algoritmo.
  - Respecto a la captura de datos, nuevamente van las dudas respecto al origen y variedad de formatos de éstos, son datos más bien estructurados por eso el formado de warehouse y recibiremos algunos no que quedaran en el Data Lake
  
3. Respecto a la arquitectura tecnológica del laboratorio, diseñada, implementada y aprobada, que permita complementar los datos capturados en el Portal con otras fuentes de datos para así medir, analizar y predecir el impacto de cursos e instancias de formación en la transformación digital y empleabilidad en MIPYMEs utilizando la metodología de medición de impacto, e implementando medidas de mitigación de riesgos en la privacidad y protección de datos personales como de los riesgos éticos de los algoritmos.
  - ¿Existe algún requerimiento tecnológico específico respecto a este punto? Respecto a este punto, deberemos tener claro además el alcance de los tipos de datos y atributos y sus respectivas fuentes de orígenes que permitan analizar y predecir el impacto de cursos e instancias de formación en la transformación digital y empleabilidad en MIPYMEs.

No hay ningún requerimiento tecnológico. Se sugiere la utilización de software existente (comercial o de código abierto) en desmedro de desarrollos propios para así facilitar el futuro soporte y mantención de los sistemas que componen la arquitectura

### **Preguntas Respecto a los Entregables:**

1. Respecto al portal web para el Laboratorio (front-end); Las consultas van referido a lo ya expuesto anteriormente, respecto a la captura de datos es tener la claridad de la fuente de origen para entender los formatos a trabajar y:

- ¿Qué datos compondrán lo referido a diagnóstico a-priori de MIPYMES?

Estamos trabajando con una herramienta que mide Madurez Digital, más otra toma de información que dice relación con las características de Mipymes y perfil del Emprendedor beneficiario.

- ¿Qué datos y atributos necesitaremos para caracterización técnico-financiera de la MIPYME, tenemos noción de todo esto y las respectivas fuentes de origen de los datos? Si
- ¿Qué criterios se utilizarán para la caracterización del emprendedor, hay criterios objetivos definidos? Si

2. Respecto al portal de captura de datos de mediciones intermedias en cursos e instancias de formación, nuevamente levanto el punto de:

- ¿Qué datos son requeridos? Con esto podremos ver los formatos a los que nos enfrentaremos, datos requeridos, contenidos, tiempos, evaluaciones intermedias, características de los formadores y del emprendedor o líder tecnológico formado
- ¿Existen formatos mínimos requeridos?

Esto es algo que tenemos que trabajar con las líneas de formación. Por ahora no tenemos un instrumento estandarizado.

3. Respecto al portal de captura de datos de diagnóstico a-posteriori MIPYME; Al igual que en uno de los puntos anteriores:

- ¿Qué datos y atributos necesitaremos para caracterización técnico-financiera de la MIPYME?
- ¿Tenemos noción de todo esto y las respectivas fuentes de origen de los datos? Si

4. Respecto a la herramienta de medición de impacto para cursos e instancias de formación en base a los datos obtenidos de los entregables anteriores:

- ¿Tenemos noción de las respectivas fuentes de origen de los datos? Si

5. Respecto al sistema de alerta temprana para MIPYMEs en proceso de formación con bajo impacto esperado en base a los datos obtenidos de los entregables anteriores:

- ¿Qué criterios son considerados bajo impacto?



El bajo impacto no es un criterio estático y fijo, sino dinámico y relativo a otras MIPYMES cursando la misma línea de formación. Las mediciones de impacto deben compararse a nivel estadístico con tests de significancia (ej: tests de detección de outliers)

6. Respecto a la herramienta de Machine Learning para el entrenamiento y uso de algoritmos predictivos;
  - ¿Existen criterios mínimos de precisión para los resultados obtenidos?

No. Estos análisis son de carácter exploratorios, por lo que no sabemos qué tipo de correlaciones hay entre variables

- ¿Existe alguna referencia mínima respecto a la sensibilidad y especificidad de los modelos propuestos? Se asumen modelos supervisados para este punto, basados en la data histórica obtenida.

Efectivamente, al inicio se construirá data para elaborar modelos supervisados, pero no se conoce de antemano la sensibilidad y especificidad de los modelos, ya que esta depende de la variedad de MIPYMES que se presenten en las líneas de formación y que tan bien cubran el universo muestral de las posibles combinaciones de variables y valores.

- ¿Cuánto es el mínimo de data histórica a recopilar?

No se sabe de antemano, dado el punto anterior. Dependerá de la variedad de MIPYMES analizadas y del grado de correlación entre variables que se encuentre. Además, se espera que la medición de impacto tome un tiempo considerable, por lo que se espera **como mínimo** un año de data histórica para comenzar a entrenar los algoritmos-

- ¿Existe algún requerimiento tecnológico específico (Python o R, por ejemplo)?

No, pero tanto Python como R son muy utilizadas en el campo y son muy bienvenidas en el proyecto.

7. Respecto a la herramienta de Business Intelligence para el análisis estadístico y visualización de datos, recomendaciones y generación de alertas tempranas;
  - ¿Existe algún requerimiento particular para el uso de una tecnología específica?

Se sugiere el uso de software existente (comercial o de código abierto) en desmedro de desarrollos propios.

- ¿Es necesario considerar licencias avanzadas para esto?

Depende de lo que se considere como licencia avanzada. Si debe considerarse la edición de reportes existentes y generación de nuevos reportes y análisis por parte de un número reducido de usuarios

- ¿Cuánto será el número de usuarios estimados que tendrá acceso a los resultados presentados en la herramienta de BI?

Los resultados deben ser accesibles por muchos usuarios, segmentados en distintas categorías, por lo que es pertinente considerar la publicación de estos dentro del front end y no necesariamente dentro de la herramienta de BI

8. Respecto a la documentación de la metodología de investigación y análisis del Laboratorio, que permita la identificación, testeo y retroalimentación de datos, indicadores e información, dependiendo de la tecnología utilizada para la confección de los modelos de Machine Learning
  - ¿Es aceptable el uso de Markdown de Colab o Jupyter Notebook para el caso de Python o el R, dependiendo del caso?

Sugerimos el uso de tecnologías más amigables e intuitivas, orientadas a usuarios que no son científicos ni programadores expertos. Por ejemplo, explorar opciones con mayor ayuda visual como analytic workflows tipo rapidminer/knime