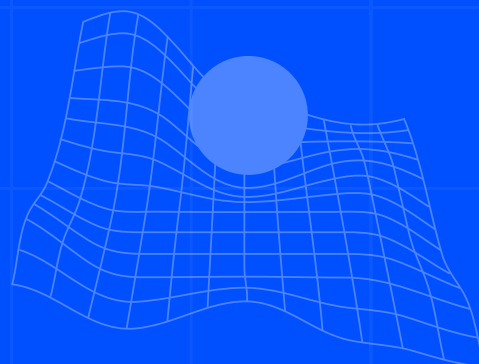
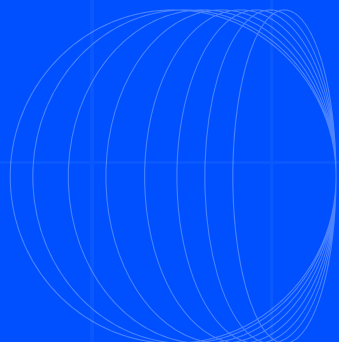




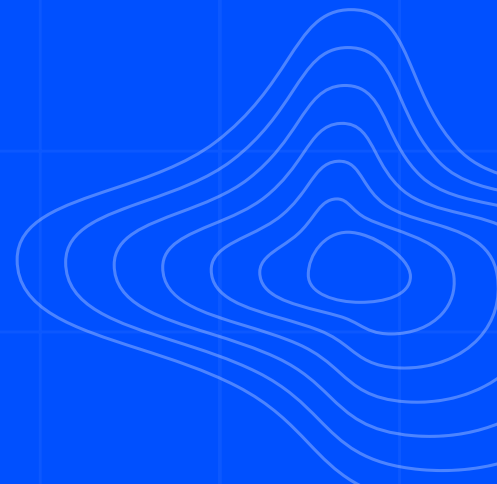
CONSEJO NACIONAL
DE CIENCIA, TECNOLOGÍA,
CONOCIMIENTO E INNOVACIÓN
PARA EL DESARROLLO



DOCUMENTO TÉCNICO

Guía metodológica para la evaluación con enfoque sistémico de Políticas CTCI

CHILE, DICIEMBRE 2025



AUTORES

María José Bravo Pizarro
Sofía Valdés Radrigán
José Zapata

CONTRAPARTE TÉCNICA

Blanca Torrico Durán
Katherine Villarroel Gatica
Catalina Terra Rosas

DISEÑO GRÁFICO

Mezcla Estudio

Este proyecto se realizó en el marco de la cooperación técnica del Banco Interamericano de Desarrollo (CH-T1332) al Consejo Nacional de Ciencia, Tecnología, Conocimiento e Innovación para el Desarrollo (Consejo CTCI).

Fecha de publicación
Enero 2026

Los Documentos de Trabajo de la Secretaría Ejecutiva del Consejo Nacional de Ciencia, Tecnología, Conocimiento e Innovación para el Desarrollo (Consejo CTCI), buscan abrir temas de discusión que permitan avanzar en el diseño consensuado de estrategias de largo plazo en estas materias, para el desarrollo de nuestro país. La siguiente guía metodológica, es parte de un proyecto de cooperación entre el Consejo Nacional de CTCI y el BID para transitar a un modelo de evaluación de políticas de CTCI con enfoque sistémico, integral, transparente y orientado al aprendizaje continuo. De esta manera, se abordan los marcos conceptuales generales y teóricos, los principales aspectos metodológicos y herramientas prácticas para su implementación, orientado a los hacedores de política pública en sus diferentes niveles.

Esta obra está bajo una Licencia Creative Commons Atribución -NoComercial- Compartir Igual 4.0 Internacional. Esta licencia significa que no se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original. Para ver una copia de esta licencia, visite:

<http://creativecommons.org/licenses/by-nc-sa/4.0/>

CÓMO CITAR ESTE DOCUMENTO

Bravo, M.J. et.al (2026) Guía metodológica para la evaluación con enfoque sistémico de Políticas CTCI. Consejo Nacional de Ciencia, Tecnología, Conocimiento e Innovación para el Desarrollo (Consejo CTCI). Santiago, Chile

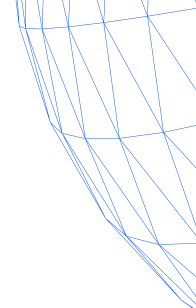
DOCUMENTO TÉCNICO

Guía metodológica para la evaluación con enfoque sistémico de Políticas CTCI

CHILE, DICIEMBRE 2025

ÍNDICE

MARCO Y VISIÓN DE LA EVALUACIÓN SISTÉMICA EN CTCI	10
Evaluación de políticas y programas en CTCI	11
Evolución hacia la evaluación sistémica	13
Principios para una evaluación sistémica	18
Alcance de la evaluación sistémica en CTCI	20
ASPECTOS METODOLÓGICOS	26
Teoría de Cambio	27
Datos en Sistemas de Monitoreo y Evaluación (M&E)	30
Metodologías de evaluación	34
ELEMENTOS CLAVE PARA UNA EVALUACIÓN CON ENFOQUE SISTÉMICO EFECTIVA	42
Teoría de Cambio anidada	45
Data de monitoreo y evaluación	50
Diseño de la evaluación sistémica	53
Desarrollo de la evaluación sistémica	57
REFLEXIONES SOBRE APLICACIÓN, SOSTENIBILIDAD Y MEJORA CONTINUA	64
REFERENCIAS	66
MATERIAL DE APOYO	69



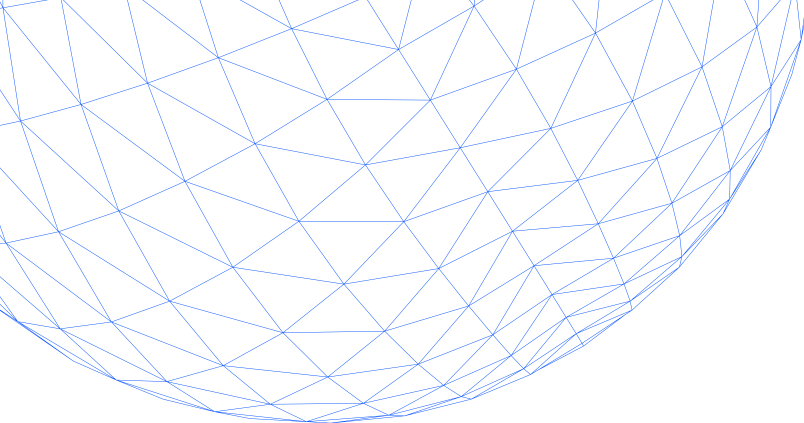
1. INTRODUCCIÓN

El presente documento se enmarca en la Hoja de Ruta - elaborada por Technopolis Group, con encargo del Consejo Nacional de Ciencia, Tecnología, Conocimiento e Innovación (CTCI) y el apoyo del Banco Interamericano de Desarrollo (BID) - para la evaluación sistémica de políticas y programas de CTCI. Dicha Hoja de Ruta propone avanzar hacia una cultura de evaluación más integral, transparente y orientada al aprendizaje, capaz de captar las particularidades de las políticas de CTCI y de articularse con los procesos de planificación y decisión pública (Technopolis, 2024).

Las evaluaciones en CTCI han estado, por lo general, centradas en programas o instrumentos aislados y/o con énfasis en resultados intermedios. Este enfoque resulta insuficiente ante los desafíos actuales, que exigen comprender trayectorias de cambio, efectos indirectos y articulación entre políticas. Frente a este contexto, se propone la evaluación sistémica como un modelo de evaluación que responde tanto a las particularidades de los desarrollos en CTCI, la evolución de los marcos y métodos de evaluación, y los grandes desafíos sociales actuales.

Transitar desde el modelo actual de evaluaciones hacia uno sistémico es un proceso que involucra muchos factores y excede los alcances de esta guía metodológica. Es por esto que se propone una evaluación con "enfoque sistémico", el cual trata de aplicar algunos de los elementos de una evaluación sistémica, en la medida que sean factibles y pertinentes, al realizar cualquier tipo de evaluación. De esta manera, en cada evaluación y con la contribución de cada evaluador, se puede ir construyendo un futuro sistema de evaluación propiamente sistémico para la CTCI.

Esta Guía consta de cuatro apartados, en el primero se abordan los marcos generales y teóricos de una evaluación sistémica de la CTCI, en el segundo se revisan aspectos metodológicos de la misma, en el tercero se desarrolla una guía práctica para la implementación de evaluaciones con enfoque sistémico, siguiendo el modelo de una lista de cotejo, mientras que en el cuarto y último se incorporan reflexiones y sugerencias sobre el proceso de instalación de una evaluación sistémica de la CTCI.



A través de estos apartados, se aborda la evaluación sistémica en distintas capas y niveles, desde lo macro (prioridades nacionales y misiones estratégicas) hasta programas e intervenciones específicas. Más que analizar cada nivel por separado, esta Guía propone mirar sus interconexiones, reconociendo la participación de actores diversos, que configuran el ecosistema CTCl, y entendiendo la evaluación como un entramado articulado.

Este material está dirigido a un espectro amplio de actores que participan en el ciclo de políticas de CTCl y que se benefician de evaluaciones:

- Tomadores de decisión: autoridades de gobierno y directivos que requieren evidencia para la toma de decisión informada para definir prioridades y asignar recursos.
- Equipos técnicos y de implementación: ministerios, agencias e instituciones ejecutoras que buscan retroalimentación para mejorar la gestión, coordinar programas y optimizar instrumentos.
- Comunidad CTCl, por ejemplo: universidades, centros de investigación, empresas, sociedad civil y oficinas de transferencia tecnológica.
- Entidades de control y fiscalización del gasto público, que requieren lineamientos claros y criterios consistentes. Instancias como la Dirección de Presupuestos (Dipres) y la futura Agencia de Calidad de las Políticas Públicas y la Productividad (ACPP).
- Público en general, ciudadanía interesada en la transparencia, rendición de cuentas y el impacto de las políticas en el desarrollo sostenible.

Su elaboración es gracias a un proceso participativo que combinó espacios de reflexión colectiva, intercambio técnico de experiencias y validación intersectorial. A través de tres talleres se generó un espacio de trabajo colaborativo entre representantes de ministerios, agencias públicas y el propio Consejo de CTCl.

El primer taller permitió sentar las bases conceptuales y metodológicas en torno a los fundamentos y criterios de una evaluación sistémica en CTCl, mediante dinámicas grupales y ejercicios de priorización colectiva. Posteriormente, el segundo taller se centró en identificar los elementos esenciales para una evaluación efectiva, analizando obstáculos y generando propuestas técnicas que nutrieron la estructura preliminar de la guía. Finalmente, en el tercer taller se revisaron y validaron los contenidos del documento, asegurando la pertinencia, aplicabilidad y utilidad de la propuesta.

El Consejo CTCl y el equipo expresan su reconocimiento a todas las instituciones y personas que participaron en estas instancias, en particular a:

Este proceso no sólo permitió recoger una amplia gama de miradas y experiencias, sino que también aseguró que los contenidos de la guía reflejen realidades y necesidades del ecosistema nacional de CTCl. Se proyecta como un instrumento dinámico, sujeto a actualización periódica bajo criterios de sostenibilidad, con apoyo de un grupo interinstitucional que refuerce su gobernanza e institucionalización en el tiempo.

Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (CTCI)

- Macarena González
- Constanza Fuentealba
- Diego Valdebenito
- Mariela Núñez-Ávila

Dirección de Presupuestos (DIPRES)

- Ernesto Laura
- David Contreras
- Alberto Sthioul

Agencia Nacional de Investigación y Desarrollo (ANID)

- Cristina Bugueño
- Valeska González

Ministerio de Economía, Fomento y Turismo

- Valentina Hernández
- Macarena Vio
- Nicolás Cerón

Ministerio de Educación (Mineduc)

- Martín Aranzaes
- Catalina Herrada
- Miguel Jerez

Corporación de Fomento de la Producción (CORFO)

- Leonor Saravia
- César Martínez

Consejo Nacional de Evaluación y Productividad (CNEP)

- Anita Zaldívar

Consejo Nacional de CTCl

- Katherine Villarroel
- Catalina Terra
- María José Menéndez

Este proceso no sólo permitió recoger una amplia gama de miradas y experiencias, sino que también aseguró que los contenidos de la guía reflejen realidades y necesidades del ecosistema nacional de CTCl. Se proyecta como un instrumento dinámico, sujeto a actualización periódica bajo criterios de sostenibilidad, con apoyo de un grupo interinstitucional que refuerce su gobernanza e institucionalización en el tiempo.

MARCO Y VISIÓN DE LA EVALUACIÓN SISTÉMICA EN CTCI



EVALUACIÓN DE POLÍTICAS Y PROGRAMAS EN CTCl

El campo de la Ciencia, Tecnología, Conocimiento e Innovación (CTCl) se desarrolla en ecosistemas complejos, caracterizados por la presencia de múltiples actores (públicos, privados, académicos y sociales), diversidad de instrumentos, interacciones no lineales y trayectorias de cambio que requieren horizontes de tiempo extendidos (Rosenberg et al., 2024; OECD, 2020). Estas características distinguen a CTCl de otros ámbitos de política pública, demandando enfoques de evaluación capaces de capturar efectos directos e indirectos, impactos tangibles e intangibles, capacidades habilitantes y adaptabilidad de las políticas (Magenta Book, HM Treasury, 2020).

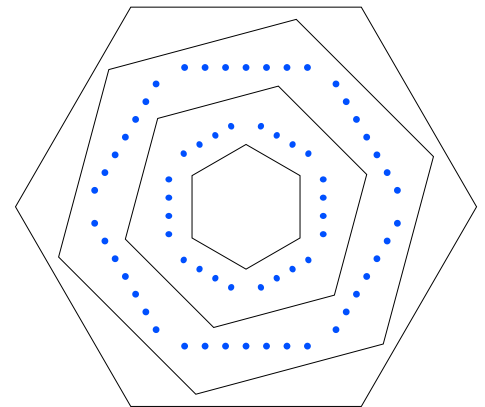
1.1 Tres generaciones de gobernanza de la CTCl

La forma de evaluar la CTCl ha tenido una trayectoria de desarrollo en conjunto con los cambios en la gobernanza de los sistemas. En este contexto, es posible identificar tres generaciones de la gobernanza CTCl (Technopolis, 2025).

1. La primera generación a partir de 1946 centrada en la investigación básica, con el informe de Vannevar Bush (1945) al presidente de EE.UU., "Science, the Endless Frontier" como manifiesto, la "delegación ciega" (Braun, 2003) de la gobernanza de la ciencia a los científicos, y confiando en el "empuje de la ciencia" para generar innovaciones y otros beneficios para la sociedad.

2. La segunda, a partir de principios de los años 60, asumió un mayor control social de la ciencia y exigió un retorno social en forma de innovación y crecimiento económico, adaptando las oportunidades científicas y tecnológicas a la "demanda", evolucionando después hacia el pensamiento de los "sistemas nacionales de innovación", que consideran la innovación como coproducida y dependiente de la participación industrial.
3. La tercera generación aborda los retos de la sociedad, como el cambio climático, las enfermedades y la pérdida de biodiversidad, e implica no sólo al sistema de I+D+i, sino también a la sociedad en general, que ayuda a decidir qué retos de la sociedad abordar, así como a poner en práctica las soluciones, dando lugar a cambios sistémicos. Esto incluye políticas basadas en retos, transiciones y "misiones" (Mazzucato, 2018; Khöler et al 2019).

Cada generación -basada en distintos propósitos-, supone desafíos particulares para la evaluación de la CTCl.



1.2 Desafíos en la Evaluación de la CTCl

El desarrollo de los sistemas CTCl, las diversas expectativas que se generan sobre ellos, y por tanto, las necesidades de abordar su evaluación, han permitido identificar una serie de desafíos específicos en este ámbito de las políticas públicas.

1. COMPLEJIDAD DEL ECOSISTEMA CTCl

En las últimas orientaciones estratégicas del sistema de CTCl en Chile se ha optado por conceptualizar el desarrollo de la CTCl y su contexto como un ecosistema. Se destaca así el avance desde una comprensión lineal de la CTCl, donde la investigación daba origen al desarrollo experimental y éste a la innovación, hacia una visión sistémica. En la mirada actual, la CTCl se genera "a través de la interacción entre los distintos actores (cada vez más diversos), en múltiples formatos, direcciones y combinaciones posibles, y no necesariamente de forma lineal en etapas" (Menéndez & Villarroel, 2024). A esta complejidad se le adiciona el entendimiento de que el sistema CTCl está inmerso en contextos económicos, sociales y ambientales, los cuales serán fundamentales para su funcionamiento, por lo que es necesario comprender todas estas dinámicas e interacciones.

2. DIVERSIDAD Y HETEROGENEIDAD DE LOS ACTORES

Para comprender el funcionamiento del ecosistema CTCl se han ido incorporando cada vez una mayor diversidad de actores, los que además de integrar más interacciones, se trata de actores distintos a los tradicionales que pertenecen al Estado, las Empresas, Instituciones de Educación Superior (IES) e Instituciones Privadas Sin Fines de Lucro (IPSFL). A éstas hoy se suman otros actores, por ejemplo los "conectores", como las Oficinas de Transferencia y Licenciamiento (OTL), Hubs de transferencia tecnológica, Incubadoras, Aceleradoras, Capital de

riesgo y los Nodos CTCl. Otros actores no tradicionales a incluir son las organizaciones de la sociedad civil, las ONG's, los medios de comunicación, clubes de ciencia, museos, escuelas, pueblos originarios, entre otros.

El desafío es el poder incorporar adecuadamente a estos actores en la evaluación CTCl, considerando su diversidad y heterogeneidad. A modo de ejemplo, la innovación no ocurre en un flujo lineal entre investigación y mercado, sino en múltiples actores, interacciones y contextos. El desafío radica en articular la diversidad de actores del ecosistema con la población, mediante criterios operativos que permitan la trazabilidad de inclusión/exclusión, mitiguen sesgos, delimiten a quién(es) observar, en línea con preguntas e indicadores para evaluar efectos esperados en el segmento del ecosistema pertinente.

3. EFECTOS DE LARGO PLAZO

La literatura de referencia indica que las inversiones en I+D tienen impactos esperados de 5 a 20 años después de la intervención. Además, es dependiente de qué tipo de desarrollo CTCl se financia o se propicia, de qué clase de investigación, la complejidad de la tecnología a desarrollar o factores de contexto. En el caso del fomento a la ciencia básica, por su propia naturaleza no se busca una aplicación directa, y sus impactos se podrán observar solo a largo plazo. Otro desafío relacionado es que cuanto mayor es el tiempo que ha pasado desde el desarrollo de la CTCl y los impactos con los que se busca relacionar, es más difícil la trazabilidad de sus causas (Rosenberg, et al., 2024).

4. BENEFICIOS INTANGIBLES

Las publicaciones académicas son un resultado documentado del desarrollo de la CTCl, pero los resultados o impactos que surgen desde estas investigaciones son complejas de registrar, cuantificar y evaluar.

Los desarrollos en CTCI generan beneficios y activos que son difíciles de medir o expresar de forma cuantitativa o monetaria. El desafío es cómo medir beneficios, como por ejemplo que la mejora en las habilidades en el saber hacer o la reputación y credibilidad de los investigadores podría hacer incidir en la obtención de mayor financiamiento de las empresas, en el fortalecimiento de las redes de colaboración de investigadores o de actores del ecosistema, entre otros.

5. ESCASA OBSERVABILIDAD

Existen efectos que se producen en lugares o procesos que son difíciles de rastrear, por ejemplo las mejoras en habilidades y capacidades científicas, tecnológicas o estratégicas, en carreras profesionales y académicas, en la movilidad de los agentes CTCI, en productividad y competitividad, en el diseño de políticas y su implementación, cambios en comportamientos, percepciones, etcétera. También por la movilidad del personal que trabaja en I+D entre instituciones, países y sectores, lo que ocasiona la pérdida de información útil para las evaluaciones.

6. EFECTOS ASIMÉTRICOS

Por cada proyecto exitoso de innovación, hay muchos casos fallidos en nuevas investigaciones, tecnologías, procesos y productos. Por lo que los métodos cuantitativos que definen promedios de resultados o impactos pueden ser problemáticos, ya que no captan bien el potencial transformador e incluso sistémico de algunos de los proyectos exitosos.

Para abordar estos desafíos no existe un método único y superior a los demás, sino que se requiere combinar métodos diversos (cuantitativos, cualitativos y basados en teoría) para captar interacciones, contribuciones múltiples, efectos indirectos y la complejidad aquí expuesta. Es por esto que ha emergido el concepto de evaluación

sistémica como la aproximación que mejor aborda hasta ahora el desafío de la evaluación en CTCI.



EVOLUCIÓN HACIA LA EVALUACIÓN SISTÉMICA

2.1 La evolución de la evaluación

El sistema de Monitoreo y Evaluación alojado en la Dirección de Presupuestos, tiene como objetivo "generar información de desempeño e introducir prácticas para mejorar la calidad del gasto, identificando en sus objetivos específicos, variables relevantes y complementarias tales como contribuir a: a) La eficiencia en la asignación de recursos, b) La eficiencia y eficacia en el uso de los recursos, c) La transparencia en la gestión de los programas y en la administración de las finanzas públicas." (Dipres, 2022).

En el libro Magenta de la Tesorería del Reino Unido, se definen como dos las principales razones para evaluar: el aprendizaje y la rendición de cuentas. El aprendizaje se trata de obtener evidencia para una mejor toma de decisiones, lo cual debería derivar en la mejora de las políticas, la reducción de riesgos, la adecuada localización de recursos, entre otros. La rendición de cuentas responde a la transparencia hacia las instituciones, actores y la ciudadanía, sobre el uso del gasto público y la efectividad de las políticas públicas. (HM Treasury, 2020).

La evolución de los marcos de política CTCI también ha marcado las tendencias con respecto a su evaluación. A nivel internacional hoy se realizan esfuerzos para relacionar la inversión en CTCI con sus contribuciones o impactos sobre grandes retos sociales, como transiciones sociotécnicas o el

concepto de "misiones". Esta tendencia ha llevado a desarrollar una evaluación sistémica, de manera de responder a la complejidad de los entornos, las múltiples interacciones entre una mayor gama de actores, la búsqueda de estos objetivos de largo plazo, e incluso indagar en los posibles efectos negativos de la investigación.

En esta misma línea, la Estrategia Nacional de Ciencia, Tecnología, Conocimiento e Innovación para el Desarrollo de Chile de 2022 se plantea una visión que asume grandes retos para el país, como lo son la plena incorporación a la sociedad del conocimiento, que el desarrollo sea sostenible, integral, que se preocupe de la preservación de la biósfera y se comprometa con la búsqueda de una sociedad inclusiva y democrática (Consejo CTCl, 2022).

Más recientemente, el Consejo Nacional de Ciencia, Tecnología, Conocimiento e Innovación para el Desarrollo, ha continuado analizando las tendencias de cambio globales y ha establecido cuatro retos, dos provenientes del avance científico e considera la revolución digital y la revolución biológica, y dos preocupaciones de contexto como son la sustentabilidad de la vida en el planeta y la crisis de la democracia y gobernanza global. También ha identificado tres movimientos que buscan dar respuesta a este escenario: a) la transición hacia economías sustentables y energías limpias; b) la demanda creciente por equidad y justicia social; y c) el uso de la ciencia y las tecnologías como factor de resiliencia, competitividad, transformación hacia lo sostenible y de geopolítica (Menéndez & Villarreal, 2024).

Se observa entonces, que la visión estratégica del país en CTCl está orientada hacia la contribución a estos grandes retos y tendencias globales, lo que sustenta la necesidad de una evaluación sistémica para responder a este escenario, además de la necesidad de la contribución a la innovación y el desarrollo.

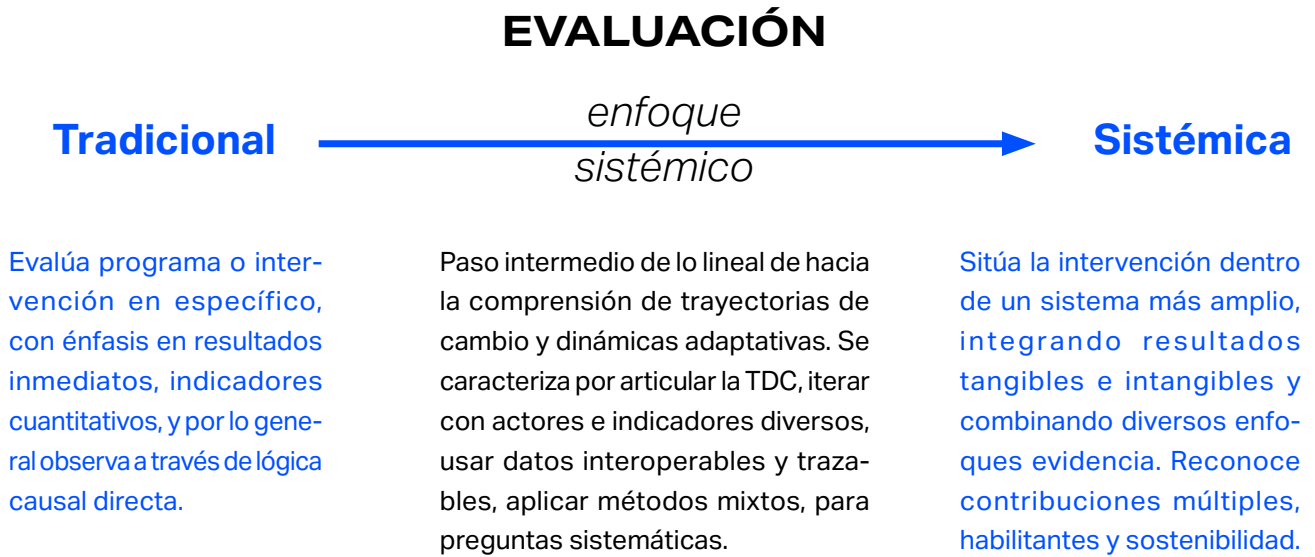
Algunas de las características de la evaluación sistémica son:

- Contar con una Teoría del Cambio (TDC) para el programa a evaluar, de manera de tener los efectos sociales previstos.
- Contar con Teorías del Cambio "anidadas" entre niveles micro, meso y sistémico de la política CTCl, es decir, que las TDC estén articuladas, donde se clarifican sus relaciones entre y dentro de cada nivel.
- Abordar la complejidad tanto del Ecosistema CTCl, como del sistema de políticas CTCl. Considerando que existen propiedades "emergentes" dada la complejidad de estos sistemas, que no pueden ser previstas, por lo que existe la necesidad de ajuste al contexto y reflexividad, es decir, el ajuste y actualización de las TDC.
- Abordar una gama mayor de actividades y actores que los que tradicionalmente se han considerado en las políticas CTCl.
- Tener un marco temporal de largo plazo y basarse en visiones rectoras, como las estrategias de CTCl, que orienten las actividades hacia la contribución a la resolución de los grandes retos y transiciones visualizadas.

2.2 Transición: De la evaluación convencional a evaluación sistémica

Avanzar hacia la evaluación sistémica en CTCl supone dejar atrás la mirada fragmentada de programas aislados para situarlos dentro de un entramado mayor de políticas, instituciones, actores y prácticas. No se trata solo de medir resultados puntuales, sino de entender cómo interactúan estrategias nacionales, capacidades organizacionales y acciones concretas.

DIAGRAMA 1: TRANSICIÓN DE LA EVALUACIÓN



Considerando que la evaluación de las políticas CTCI es una parte del sistema de políticas CTCI, el que a su vez, es un componente del ecosistema de la CTCI, se hace necesario especificar que una "concepción sistémica" puede aplicarse a estos tres ámbitos, que si bien están relacionados, es pertinente distinguirlos para comprender el alcance de lo que sería un "enfoque sistémico".

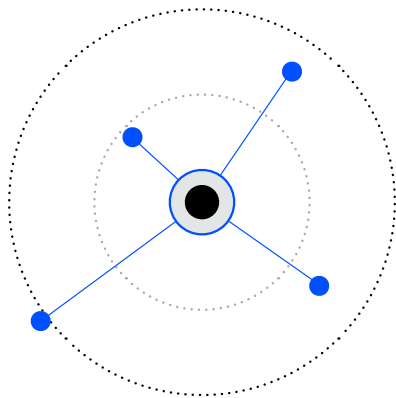


DIAGRAMA 2: ECOSISTEMA CTCI



FUENTE: ELABORACIÓN PROPIA

TABLA 1: DIMENSIÓN DEL ECOSISTEMA CTCl

Dimensión	Concepción Sistémica
Ecosistema CTCl	Los actores CTCl definen propósitos globales compartidos (misiones/transiciones), se coordinan y colaboran.
Políticas CTCl	Sistema conscientemente articulado de regulaciones, programas, incentivos y acciones con propósitos comunes y en permanente ajuste con los cambios del ecosistema.
Sistema Evaluación CTCl	Sistema de monitoreo y evaluación conscientemente articulado para evaluar la política CTCl en sus distintos niveles (macro, meso y micro), como parte del ecosistema CTCl, en función del cumplimiento tanto de los propósitos globales (ej desarrollo sostenible) como de los objetivos de las políticas, instituciones y programas.

El logro de una concepción sistémica tanto para el Ecosistema CTCl, las políticas CTCl y para su sistema de evaluación y monitoreo es una transformación por cumplir, la cual implicará voluntades, tiempo y recursos. Sin embargo, para avanzar en ese camino, se propone la adopción de un “enfoque sistémico”, el cual consiste en ir construyendo esa concepción sistémica por medio de aplicar conceptos y acciones en cada contexto, que permitan avanzar desde el estado actual hacia el que se busca. En función de la evaluación, el enfoque sistémico puede ser aplicado a cualquier nivel de evaluación, macro, meso y/o micro, donde en cada parte del proceso se busque avanzar hacia una concepción sistémica. Sin pretender exhaustividad, son elementos parte del enfoque sistémico aplicado a la evaluación CTCl:

- Analizar cualquier programa o componente como parte del ecosistema CTCl, apuntando a sus interacciones complejas (tangibles e intangibles) y buscando las interconexiones entre cada elemento.
- Considerar marcos temporales adecuados, tanto en los antecedentes como en las proyecciones, para la observación de resultados e impactos en el área CTCl.
- Realizar acciones que busquen la información, comunicación, involucramiento y coordinación de los actores CTCl.
- Buscar la convergencia de las miradas de los actores CTCl en torno a propósitos globales comunes.
- Buscar la articulación del sistema de monitoreo y evaluación de la CTCl, en todas las partes del proceso, de manera que cada evaluación e indicador contribuya a ampliar el conocimiento tanto del conjunto de políticas CTCl en sus interacciones y del ecosistema CTCl.
- Complementar distintas metodologías de evaluación, con sentido crítico e innovador, para ir avanzando hacia modelos evaluativos sistémicos, los cuales permitan abordar la complejidad del ecosistema CTCl y de las transformaciones que buscan los propósitos globales comunes.

2.3 ¿Por qué avanzar hacia evaluación sistémica?

Además de las razones ya presentadas que justifican avanzar hacia una evaluación sistémica, también existen debilidades detectadas y demandas de los actores locales vinculados con las evaluaciones, que podrían abordarse en una transición hacia un enfoque sistémico.

Necesidad de articulación en la oferta política y programática CTCl

Se identifica una insuficiente articulación de la oferta programática de CTCl, incluso dentro de una misma institución. Esto puede estar generando superposiciones no buscadas entre los programas, instrumentos y llamados, lo que supondría interacciones tanto positivas (sinergias) como negativas, que podrían no ser potenciadas en el primer caso o mitigadas en el segundo. Una de las preocupaciones principales relacionadas es la posible ineficiencia que podría generarse por este motivo.

Necesidad de una integración de las bases de datos CTCl

Se requiere ampliar la información sobre los resultados de los programas e integrar los datos del sistema mediante métricas y diccionarios de datos comunes, asegurando interoperabilidad técnica y semántica. Esto implica usar estándares compartidos (formatos, metadatos, definiciones) y mecanismos de vinculación interoperables que habiliten cruces confiables entre fuentes, por ejemplo, identificadores persistentes o reglas de correspondencia entre registros. Así los datos de monitoreo y evaluaciones anteriores serán (re) utilizables para evaluaciones futuras, facilitando el cumplimiento de objetivos y reduciendo los costos de levantamiento de información.

Avanzar hacia metodologías que respondan adecuadamente a las políticas CTCl

Se diagnostica que la metodología del marco lógico no logra captar toda la complejidad de los programas e instrumentos CTCl, ya que se enfoca principalmente en aspectos financieros y de eficiencia, no contemplando adecuadamente el impacto a largo plazo, la generación de capacidades o la articulación de los programas dentro del ecosistema CTCl. Además existen contextos locales o regionales que tampoco logran ser considerados adecuadamente. Es por esto que se sugiere adaptar el marco metodológico, complementándolo con enfoques más flexibles y holísticos, como la Teoría del Cambio, que permitan evaluar la complejidad y el impacto real de las intervenciones en el ecosistema.

Seleccionar los programas a evaluar según un marco estratégico e involucrando a los actores relacionados

Actualmente, el proceso de selección de programas a evaluar no tiene criterios claros y reconocidos por todos los actores, lo que dificulta la garantía de que las evaluaciones respondan a prioridades estratégicas. Por otro lado, se percibe un bajo involucramiento de los actores incumbentes en las evaluaciones en estas definiciones, lo que hace perder perspectivas y retroalimentación a estas decisiones. Este bajo involucramiento reduce la utilidad práctica de las evaluaciones.

Es por esto que se propone avanzar hacia la creación de instancias participativas y un marco metodológico consensuado que permita articular mejor a los diversos actores, generando de esta forma una mayor eficiencia en el uso de los recursos para las evaluaciones.

Posibilidad de prever efectos sistémicos de modificaciones a los programas en implementación del sistema CTCl

La articulación explícita y por medio de teorías del cambio anidadas que traería la posibilidad de conocer las interacciones esperadas entre programas, permitiría considerar el cambio en esas interacciones y el efecto sistémico que provocaría el actualizar, modificar, escalar o eliminar algún programa CTCl que se estuviera implementando dentro del sistema.



PRINCIPIOS PARA UNA EVALUACIÓN SISTÉMICA

Para que la evaluación en CTCl cumpla con sus fines de aprendizaje y rendición de cuentas, debe concebirse como un proceso sistémico, esto es: planificado desde el diseño de la política, ajustado en su alcance y métodos según el tipo de programa, y orientado a producir evidencia útil en los momentos críticos del ciclo de gestión. En este marco, se proponen un conjunto de principios de buena práctica operativa, que permiten guiar el diseño y ejecución de las evaluaciones, asegurar su calidad y aumentar sus posibilidades de uso en la toma de decisiones. Estos principios no son prescripciones rígidas, sino orientaciones que pueden adaptarse a distintos contextos y escalas de intervención. En conjunto, buscan fortalecer la capacidad del sistema para aprender, mejorar y rendir cuentas de manera transparente.

UTILIDAD Y USABILIDAD

Exige que las evaluaciones respondan a preguntas de decisión concretas y entreguen productos utilizables en los momentos oportunos. En la práctica, esto supone mapear con precisión los hitos del

ciclo de gestión (por ejemplo, los presupuestarios, actualizaciones de programas, cierres de convocatorias) y calendarizar entregables contra esos hitos. Otra buena práctica operativa es entregar la evidencia en el formato que cada actor necesita, cuando la necesita.

- **Autoridades:** brief ejecutivo con hallazgos y alternativas de decisión
- **Gestión:** informe ampliado, con análisis de desempeño, identificación de cuellos de botella, próximos pasos, entre otros contenidos relevantes.
- **Técnico:** anexos metodológicos, abordando métodos, supuestos, validez.

INTEGRACIÓN EN GOBERNANZA

La evaluación aporta valor y es útil en la medida que esté integrada en la gobernanza de los ciclos de política pública. De este modo, las recomendaciones emanadas de estudios no quedan "en el aire", sino que se incorporan en los procesos formales de revisión de programas y su financiamiento. Algunos mecanismos de articulación para favorecer el uso efectivo de la evidencia pueden ser:

- Comités la conducción de la evaluación, como instancias que acompañan el diseño y las principales etapas de ejecución, asegurando desde el inicio el acceso a datos, calidad de la información, así como la definición de mecanismos para la adopción de recomendaciones.
- Revisión de pares: proceso que otorga legitimidad y robustez técnica a las evaluaciones previo a su publicación, fortaleciendo la confianza en sus hallazgos.
- Vinculación con el ciclo presupuestario (DIPRES), usando resultados de la evaluación en etapas clave, tales como en el diseño o reformulación de programas, en su monitoreo, y en la discusión presupuestaria.

En el caso de programas orientados a transiciones socio-técnicas -tales como los en materia en CTCl-, esta integración debiera ser multinivel y multiactor, conectando la gestión para avanzar en los cambios esperados.

PROPORCIONALIDAD

Ajustar la profundidad de la evaluación, su alcance y métodos de forma proporcional al riesgo, novedad, tamaño y necesidades de información de cada programa. Esto permite optimizar recursos y esfuerzos donde se genere mayor valor público con la evidencia producida. Lo primero es transparentar que no todas las intervenciones requieren del mismo esfuerzo evaluativo, y que por ende, aplicar este principio puede verse así:

- En pilotos o de bajo riesgo, evaluaciones más ligeras: por ejemplo, seguimiento de indicadores, revisión de procesos, con informes breves que permitan tomar decisiones rápidas de ajuste.
- En programas de mediana escala, evaluaciones intermedias, combinando, por ejemplo, técnicas de seguimiento con estudios de casos, consultas a actores clave.
- En políticas estratégicas y de alto costo, evaluaciones intensivas, a través de métodos complejos, complementadas con revisión de pares y/o análisis de sensibilidad.

ARTICULACIÓN

Los instrumentos en materia de CTCl suelen operar en la lógica de portafolios, donde las intervenciones se solapan, complementan y/o encadenan. Por tanto, evaluar cada uno de manera aislada o “en silo” puede ocultar sinergias, duplicidades y restricciones vinculantes de capacidad, coordinación, secuencia, normativa e información. Así, este

principio promueve evaluaciones que consideren la interacción entre programas y que coordinen a los distintos actores involucrados en su diseño y ejecución. Para abordar la articulación entre instrumentos se puede realizar:

- Evaluaciones de portafolio: análisis conjunto de programas complementarios.
- Teorías de cambio anidadas para representar de forma explícita las conexiones y límites de cada instrumento en relación con otros.

Articular diversos actores en torno a la evaluación puede implementarse a través de:

- Coordinaciones interinstitucionales, como instancias de trabajo conjunto entre ministerios, agencias y ejecutores que facilitan el intercambio de información y la planificación de evaluaciones.
- Gobernanzas multinivel: mecanismos de articulación que integran distintos niveles de decisión (nacional, regional, local) y actores públicos, privados y sociales.

SOLIDEZ Y RIGUROSIDAD

Busca que las conclusiones sean válidas, confiables y útiles para la toma de decisiones, evitando interpretaciones apresuradas o basadas en evidencia insuficiente. Se expresa en la adecuada correspondencia entre las preguntas planteadas y los métodos seleccionados, en la existencia de mecanismos de garantía de calidad, así como en:

- Tamaño muestral y alcance claros, definiendo desde el inicio el poder estadístico, el alcance de las inferencias y las limitaciones de los datos.
- Transparencia en supuestos: explicitar las hipótesis del diseño, los márgenes de error y la sensibilidad de los resultados frente a cambios en los supuestos.

- Triangulación de métodos y fuentes, combinando enfoques cuantitativos y cualitativos para fortalecer la robustez de las conclusiones.

En el caso de las políticas de I+D+i, que suelen tener efectos rezagados y contrafactuales complejos, resulta importante combinar enfoques y aplicar triangulación sistemática. La evidencia cualitativa no cumple solo un rol ilustrativo; es un insumo central para entender mecanismos, contextos y condiciones habilitantes que los datos cuantitativos no capturan por sí solos. Su validez descansa en criterios de diseño tan exigentes como los de los métodos estadísticos: definición explícita de la muestra y de sus límites, guías de levantamiento ancladas en la Teoría de Cambio, procedimientos sistemáticos de codificación y análisis, y trazabilidad entre citas, hallazgos y conclusiones. Un estudio cualitativo bien diseñado permite, por ejemplo, someter a prueba supuestos críticos, identificar supuestos, explicar heterogeneidad de resultados entre territorios, tipos de actor o poblaciones.

CULTURA DE MEJORA CONTINUA

El propósito último de la evaluación no es solo rendir cuentas, sino también mejorar los programas y fortalecer la capacidad institucional para futuras intervenciones. Este principio destaca la necesidad de planificar el uso de la evidencia, más allá de su producción, generando aprendizajes que se traduzcan en cambios concretos en el diseño, la gestión y la formulación de nuevas políticas. La mejora continua requiere tanto de mecanismos de diseminación y retroalimentación como de la instalación de capacidades en los equipos públicos y de implementación. Algunas buenas prácticas son:

- Retroalimentación sistemática: realizar revisiones posteriores a cada ciclo, como sesiones de *after-action review*, para ajustar programas e instrumentos a partir de los hallazgos.

- Institucionalización de la mejora: integrar procesos y aprendizajes de evaluación en normas, lineamientos y rutinas organizacionales que trascienden a las personas.

De esta manera, la evaluación se convierte en un motor de aprendizaje organizacional, siendo condición necesaria que exista apertura al cambio y que las instituciones promuevan la cultura de mejora continua.



ALCANCE DE LA EVALUACIÓN SISTÉMICA EN CTCI

Al hablar de alcance de la evaluación en CTCI, no se trata únicamente de definir qué resultados medir, sino de precisar en qué escala del sistema se realiza el análisis (niveles), en qué momento del ciclo de intervención se aplica y qué dimensiones se privilegian (focos de análisis).

La combinación de escalas, temporalidades y dimensiones permite responder a los desafíos de complejidad de los ecosistemas de innovación y asegura que la evaluación aporte valor no sólo en términos de rendición de cuentas, sino también como insumo para la mejora continua, el aprendizaje institucional y la legitimidad del sistema en su conjunto.

4.1 Niveles de evaluación

En el marco de la evaluación sistémica, los niveles se entienden como las distintas escalas de decisión y acción en las que una política, programa o proyecto puede ser analizado. Cada nivel responde a lógicas propias de diseño, gestión y uso de resultados, involucra actores diferenciados y plantea preguntas específicas.

Lo relevante es que no se trata de compartimentos aislados, sino de planos interconectados: lo que ocurre en el nivel de proyecto influye en los portafolios programáticos, y estos, a su vez, condicionan la efectividad de las estrategias y políticas a nivel nacional.

TABLA 2: NIVELES DE EVALUACIÓN Y SU ALCANCE SISTÉMICO

Nivel	¿Qué evalúa?	Actores involucrados	Alcance sistémico
Político - Estratégico MACRO	Políticas nacionales de CTCI, estrategias sectoriales, misiones o agendas nacionales.	<ul style="list-style-type: none"> Consejo CTCI, Ministerios (ej: Ciencia, Economía, Educación) Entidades de planificación y presupuesto (ej: Hacienda, Dipres) Coordinación interministerial para alinear prioridades y marcos de evaluación 	Propiamente "sistémico": orienta la dirección del sistema en su conjunto y alinea prioridades.
Programático y/o portafolio de instrumentos MESO	Programas, fondos, concursos o conjuntos de instrumentos vinculados a objetivos estratégicos	<ul style="list-style-type: none"> Agencias públicas ejecutoras de programas, organismos de financiamiento Consejos sectoriales Unidades de monitoreo y evaluación de las agencias 	Evaluación de grupo de programas según su relación sistémica, considerando la complementariedad y coherencia de instrumentos como parte de un portafolio.
Instrumento, de proyecto MICRO	Proyectos particulares, pilotos, intervenciones locales o institucionales	<ul style="list-style-type: none"> Instituciones beneficiarias (ej: universidades, centros de I+D, empresas). Equipos ejecutores de proyectos. Evaluadores externos 	Evaluación con enfoque sistémico, al aplicar las dimensiones de esta guía (ya mencionadas arriba)

FUENTE: ELABORACIÓN PROPIA

4.2 Focos de análisis

La evaluación de políticas y programas en ciencia, tecnología e innovación no puede concebirse como un ejercicio único o estático. Su verdadero potencial radica en desplegarse a lo largo de todo el ciclo de intervención, con distintos propósitos y metodologías que se complementan entre sí. Tanto la literatura internacional como el sistema nacional de evaluación de programas de DIPRES coinciden en distinguir momentos clave: diseño (ex ante), proceso (ex dure), resultados (ex post) e impacto.

DISEÑO	→	PROCESOS	→	RESULTADOS	→	IMPACTO
Evalúa la pertinencia, coherencia y factibilidad de la intervención antes de implementarla. Permite anticipar brechas de capacidades, complementariedad de instrumentos y riesgos sistémicos.		Durante la ejecución, examina la calidad de la gestión, la gobernanza y la coordinación interinstitucional. Favorece el aprendizaje y los ajustes en contextos dinámicos.		Al cierre, mide outputs inmediatos y outcomes de corto y mediano plazo. Informa sobre retornos tempranos y distribución territorial o sectorial de beneficios. EPG en DIPRES.		Busca capturar transformaciones de largo plazo (productivas, institucionales o sociale[s]), reconociendo que en CTCl los efectos son intangibles, emergentes y resultado de múltiples contribuciones.

Así la evaluación en CTCl debe entenderse como un ejercicio multinivel y dinámico; estos planos no son compartimentos aislados, sino escalas interconectadas. Al ser la evaluación un proceso temporal y cíclico, acompaña a las intervenciones desde su diseño hasta la medición de resultados e impactos, siendo un proceso continuo, integral y adaptativo, que combina distintas escalas, tiempos y dimensiones de análisis.

4.3 Prácticas de evaluación con orientación sistémica

La orientación sistémica asume que las intervenciones operan dentro de un entramado de políticas, fondos, instituciones y actores que interactúan entre sí. Por ello, además de evaluar programas aislados, se busca entender cómo el conjunto (políticas, portafolios e iniciativas) se

alinea, se coordina y produce contribuciones hacia misiones, transiciones y objetivos nacionales de largo plazo. En esta sección se presentan distintas prácticas o tipos de evaluación con enfoque sistémico, mostrando cómo este tipo de evaluaciones puede adaptarse a distintas escalas.

Evaluación de alto nivel

Evaluaciones globales y estratégicas del sistema, usualmente promovidas por autoridades centrales (congreso, presidencia) y encargadas a expertos reconocidos ampliamente. No buscan una atribución causal directa de los resultados o impactos de la política CTCl, sino diagnosticar sus fortalezas y debilidades enfocándose en una participación amplia de los actores y realización de consultas. Normalmente su propósito es generar un rediseño institucional o incorporar cambios importantes.

Sus principales requerimientos son el consenso e impulso del sistema político, para su realización, legitimidad e implementación de los resultados. Sus ventajas son propiciar cambios importantes en el sistema y facilitar la implementación de las recomendaciones al contar con impulso político y propiciar la participación. Como desventajas es que no establece relaciones causales ni genera cifras que sustenten tales relaciones.

Un ejemplo internacional es la Nurse Review de 2015 del Reino Unido. Se trató de una revisión estructural del sistema de Consejos de Investigación británicos, encargada por el gobierno, con amplia consulta pública, entrevistas a expertos y análisis documental para identificar fallas de coordinación y redundancias institucionales. Se enfocó en revisar la arquitectura institucional del sistema y proponer un rediseño estratégico. Su principal resultado fue recomendar la creación de UKRI como agencia unificadora de la CTCI en Reino Unido.

En Chile, hay precedentes de este tipo de evaluación: se le encarga al Consejo Nacional para la Innovación para la Competitividad (CNIC) la elaboración de la Estrategia Nacional de Innovación (2007-2008), la cual contempló una evaluación y propone el primer diseño institucional de CTCI en Chile. Luego, durante el segundo gobierno de la Presidenta Bachelet se encomienda un nuevo estudio al mismo Consejo Nacional que da origen a la Comisión de Ciencia para el Desarrollo, la cual dentro de sus conclusiones propone el actual diseño institucional.

Evaluación de fondos

Miden la implementación, resultados e impactos de fondos, que refieren a un conjunto de programas (con diferentes temáticas, focos e instrumentos) cuyo financiamiento tiene un origen común, y apuntan a objetivos también comunes. Se generan evaluaciones tanto de los programas como del

fondo en su conjunto. Dependiendo del nivel que se haga de la evaluación, permite establecer causalidad y entrega datos cuantitativos.

Sus requerimientos de datos son según las metodologías que se apliquen. Otros requerimientos generales son recursos, tiempos y competencias específicas. Sus ventajas son que busca establecer causalidades y entregar datos cuantitativos de respaldo. También logra ser robusta cuando se integran varios métodos de investigación. Sus desventajas son el requerimiento de datos según la metodología que se utilice.

Un ejemplo internacional es la evaluación de Programas de innovación empresarial ANII de Uruguay. Se evaluaron un conjunto de instrumentos de innovación empresarial. Se hizo un diseño cuasi-experimental, utilizando "propensity score matching", diferencias en diferencias, para estimar el impacto de los instrumentos de ANII sobre el desempeño empresarial. Se combinaron encuestas ex post, entrevistas semiestructuradas y análisis de trayectoria de beneficiarios. Entre sus resultados estuvo el identificar mejoras en desempeño económico e innovación (Bukstein et. al., 2020).

Evaluación de subsistemas: sectorial, disciplinar, por dimensión o misión

Evaluaciones que analizan, como un sistema integrado, el desempeño, la coherencia y la contribución conjunta de políticas, programas e instrumentos que inciden sobre un mismo sector, misión o dimensión transversal. Su unidad de análisis es el subsistema específico. A diferencia de la evaluación de un programa aislado, buscan detectar nodos críticos y mecanismos habilitadores, medir resultados e impactos visibles e intangibles, y conducir la convergencia de políticas y actores en distintas escalas hacia misiones u objetivos comunes. En su conjunto, considera actores públicos y

privados, instrumentos de oferta y demanda, cadenas de valor, arreglos de gobernanza y diferencias territoriales.

En la práctica internacional, varios países han desarrollado ejercicios de evaluación de la calidad de la investigación y de su impacto por áreas temáticas o disciplinas, sin asociarlos directamente a instrumentos específicos, sino a trayectorias históricas de política. Destacan el Research Excellence Framework (REF) en UK, que concentra cerca del 60% de la asignación presupuestaria en investigación por períodos de siete años; las evaluaciones disciplinarias en Noruega, orientadas a detectar

fallas sistémicas y definir apoyos focalizados (como el fortalecimiento de carreras intermedias); y las evaluaciones periódicas de institutos de investigación en Canadá, con énfasis en demostrar impacto académico y socioeconómico (Technopolis, 2025).

Estos ejercicios, de alto costo y complejidad, combinan indicadores cuantitativos (como bibliometría), evidencia cualitativa (casos de impacto) y revisión por pares, ofreciendo una referencia relevante para Chile sobre cómo articular evaluaciones sectoriales y disciplinares con la gobernanza y financiamiento de la investigación.

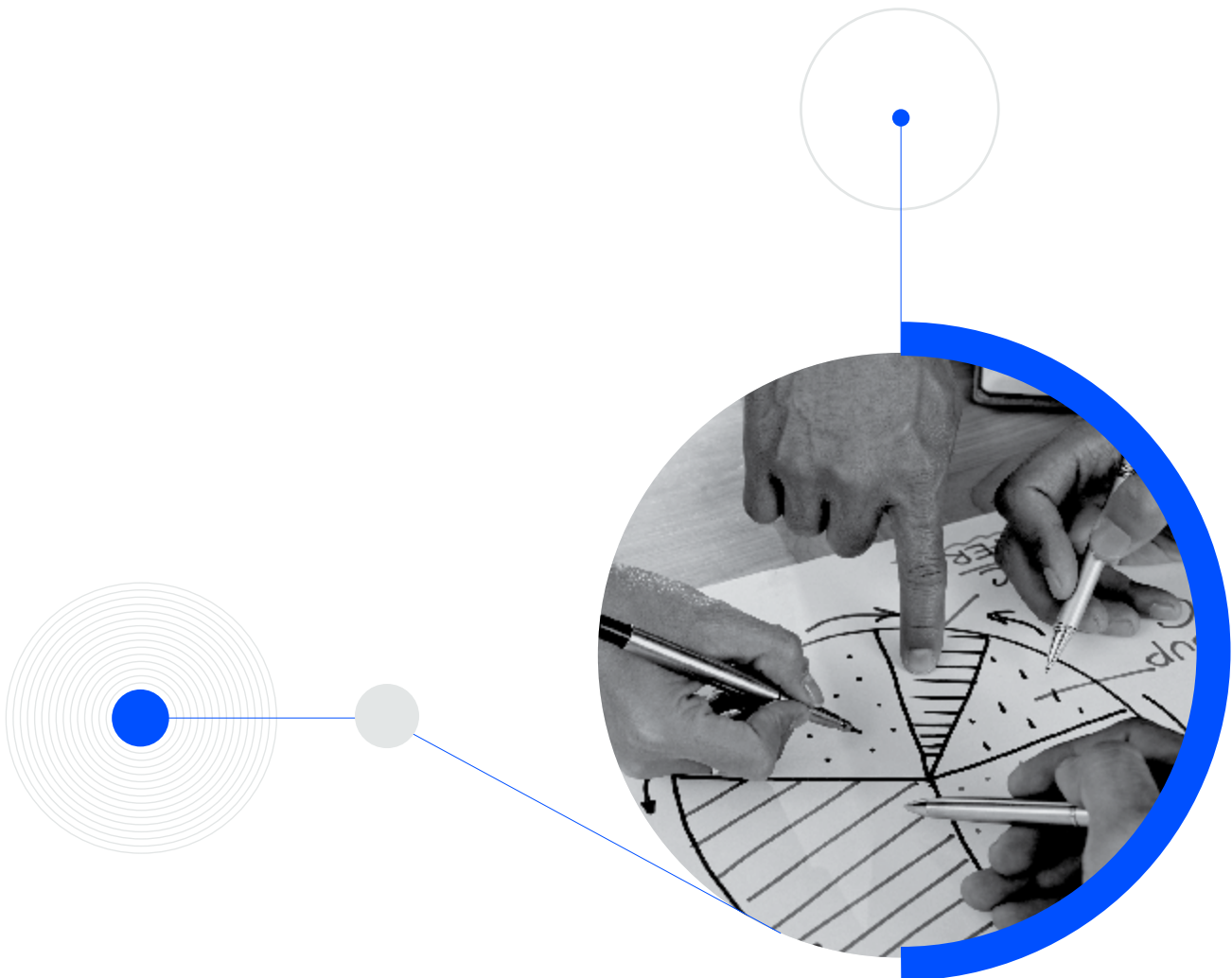


TABLA 3: EVALUACIONES DE SUBSISTEMAS EN EL CONTEXTO CTCl

Tipo	Descripción	Ejemplo
Sectorial	Analiza, como sistema integrado, políticas e instrumentos que impactan un sector económico (ej: energía, agro, salud). Clasificación: CIU para actividad económica	Evaluación sectorial (agro) del sistema de innovación en fruticultura de exportación, focalizado en uva de mesa y arándanos. Se analizarían programas de I+D en variedades resistentes a cambio climático (ANID, FIA), instrumentos de CORFO, regulaciones de SAG y protocolos de inocuidad de mercados internacionales.
Área y disciplina de conocimiento	Evalúa el desempeño de un campo específico de conocimiento, considerando sus capacidades de I+D, formación, resultados e impactos. Clasificación: <ul style="list-style-type: none"> • Manual de Frascati (OCDE), si el foco es en I+D • CINE (UNESCO), si el foco es formación 	Una evaluación disciplinar en Biotecnología en salud podría analizar: producción científica, formación avanzada, mecanismos de transferencia tecnológica, valorar la incidencia en políticas públicas de salud.
Dimensión sistémica	Evalúa una función sistémica de las diferentes etapas dentro del ciclo I+D+i y mecanismos para impulsar actividades de CTCl. Por ejemplo, transferencia tecnológica, capital humano avanzado, I+D empresarial, infraestructura.	Evaluación de la transferencia tecnológica en universidades chilenas; desempeño de OTLs, contratos de I+D con empresas, ingresos por licencias y creación de spin-offs.
Misión	Analiza la contribución conjunta de políticas, programas e instrumentos a un objetivo de gran escala, con metas medibles y horizonte definido.	Descarbonización industrial: evaluación de pilotos de electrificación, incentivos de demanda, regulación de emisiones y formación técnica asociada.

FUENTE: ELABORACIÓN PROPIA

ASPECTOS METODOLÓGICOS



TEORÍA DE CAMBIO

La Teoría de Cambio (TDC) es una descripción y representación rigurosa de cómo y por qué se espera que una intervención produzca resultados e impactos en un contexto determinado. A partir de los impactos de largo plazo, conecta insumos, actividades y productos con resultados intermedios, hace visibles los mecanismos causales, explicita supuestos y riesgos, y reconoce factores externos que condicionan las trayectorias.

1.1 Transición de herramientas metodológicas para la evaluación con enfoque sistémico

La Matriz de Marco Lógico (MML) y la cadena que va desde los insumos hasta resultados aportan la secuencia y el lenguaje común para ubicar indicadores. La MML funciona como un contrato de gestión y medición: define objetivos por nivel, indicadores, líneas base, metas, fuentes, supuestos.

La Teoría del Cambio (TDC) no reemplaza esta base: la amplía al explicar cómo y por qué se espera el cambio, incorporando rutas múltiples o condicionales, dependencias entre actores y calidad/vacíos de evidencia. Considerando que los procesos de desarrollo científico - tecnológico se caracterizan por sus efectos emergentes, bucles de retroalimentación, y alta interdependencia, la TDC permite formular hipótesis evaluables, gestionar la incertidumbre con revisiones por etapas y mantener rendición de cuentas. Se recomienda trabajar con marcos anidados (instrumento, programa, portafolio o misión).

Esta transición implica cuatro cambios de enfoque (conceptuales y operativos), que permiten a la

TDC sostener evaluaciones más robustas, útiles y adaptativas:

1. **De secuencias internas a causalidad contextualizada:** Ya no basta con describir una secuencia interna de insumos, actividades y productos: la evaluación sistémica requiere explicitar el contexto como parte de la causalidad. Supuestos críticos, riesgos y factores externos deben identificarse, monitorearse y gatillar ajustes si cambian.
2. **De instrumentos aislados a marcos anidados y portafolios:** La TDC vincula programas con metas superiores (misiones, estrategias) y orienta la gestión de portafolios dinámicos con revisiones periódicas para la incorporación de aprendizajes y la contribución conjunta.
3. **De productos únicos a rutas múltiples e indicadores de capacidades:** Además de productos y resultados tangibles, se reconocen condiciones habilitantes (capacidades, coordinación, regulación y redes) como resultados intermedios.
4. **De gestión por cumplimiento a gobernanza adaptativa:** La gobernanza se diseña en función de la TDC: estructuras, mandatos y reglas que habiliten decisiones compartidas, aprendizaje y ajustes informados a lo largo del ciclo. Para ordenar roles y responsabilidades de esta gobernanza, la guía adopta la matriz RACI¹.

Así la Teoría de Cambio es mucho más que un diagrama de "cajas y flechas", opera como hilo conductor que articula diseño, monitoreo con sus respectivos indicadores y evaluación en el marco del Ciclo de Aprendizaje: ordena la lógica, guía preguntas y métodos, y se actualiza con la evidencia para mejorar decisiones y la contribución de la intervención al cambio sistémico.

¹ RACI: Responsable, Aprobador, Consultado e Informado

1.2 Elementos de la Teoría de Cambio

Diagnóstico del problema o necesidades

Descripción del fenómeno público que origina la intervención y de las condiciones en que se manifiesta. Debe identificar las causas inmediatas y subyacentes, precisar su magnitud, tendencias, inequidades transversales (territoriales, de género, tipo de actor, entre otras), así como los costos de no actuar frente al escenario base. Define la población afectada y la población objetivo, estableciendo criterios de focalización y consideraciones de equidad. Sustenta el diagnóstico en evidencia disponible y explícita vacíos de información.

Propósito

Objetivo general o resultado principal que se espera obtener en la población, una vez que se haya ejecutada la intervención, identificando su nivel y

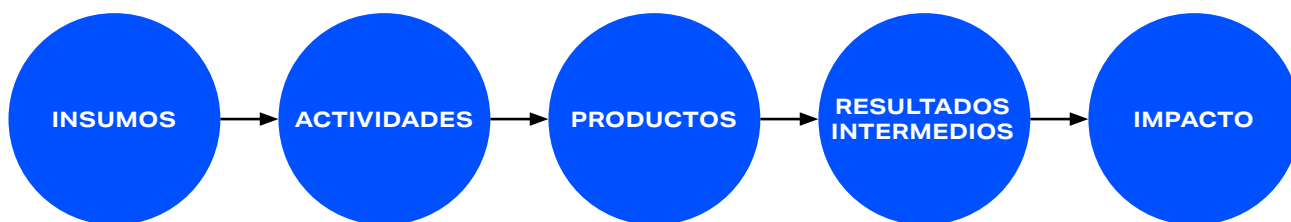
alcance (programa, instrumento, portafolio/misión). Se puede entender también como la razón que da origen a la existencia de la intervención.

Impacto

Entendido como objetivos o metas de largo plazo y/o resultados finales a nivel de sistema/beneficiarios. Generalmente se definen con criterios de éxito y suficiencia, incorporan consideraciones de equidad y se miden en ventanas temporales acordes a los rezagos del sector. Cuando el cambio dependa de múltiples intervenciones, la evaluación con enfoque sistémico se centrará en la contribución (y no en atribución exclusiva) del programa al cambio observado.

Cadena de resultados

Representación operativa de la Teoría de Cambio que ordena, en una secuencia causal trazable, la relación entre:



→ **Insumos:** recursos que el programa moviliza o controla para operar (financieros, humanos, tiempo, datos, capacidades, normativa, alianzas). Son las condiciones de partida para ejecutar la intervención.

→ **Actividades:** acciones realizadas con esos insumos para producir entregables (por ejemplo, convocatorias y selección, entrega de financiamiento, adquisición o desarrollo de soluciones, asistencia técnica, formación, coordinación). En esquemas sintéticos puede consolidarse con insumos.

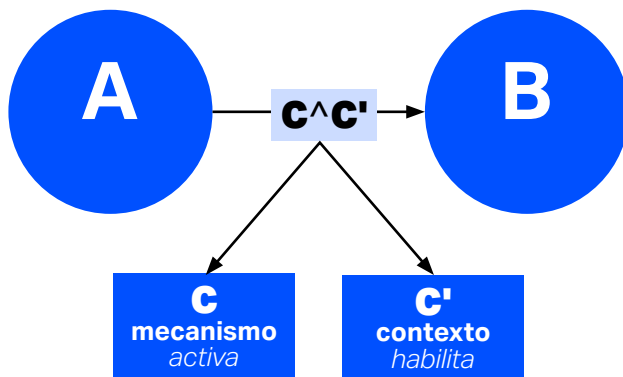
→ **Productos:** entregables inmediatos, verificables y bajo control del programa, expresados como trabajo terminado. Surgen de la combinación de insumos y actividades y pueden materializarse como bienes y/o servicios.

→ **Resultados intermedios (outcomes):** son señales tempranas de cambios observables en prácticas, capacidades, comportamientos o relaciones. Anticipan resultados finales/impacto y dependen de supuestos y condiciones habilitantes.

→ **Impacto:** transformaciones de mayor escala y plazo, atribuibles al conjunto del portafolio del ecosistema. La TDC da cuenta de contribución del programa, no de la atribución exclusiva.

La cadena explícita vías causales: secuencias ordenadas de cambios que conectan insumos y actividades con productos, resultados e impacto, reconociendo rutas alternativas que pueden conducir al mismo objetivo.

Para cada enlace $A \rightarrow B$ se registra la lógica causal: si ocurre A, entonces esperamos B, porque opera el mecanismo C, bajo las condiciones C'.



- Supuestos: Condiciones del entorno, fuera del control directo del programa, que deben cumplirse para que los enlaces causales ($A \rightarrow B$) funcionen según lo previsto. Se formulan como hipótesis verificables y se priorizan las críticas.
- Factores externos (contexto): Características del sistema (reglas, mercado, tecnología, capacidades institucionales, cultura, territorio) que influyen en la cadena, sin ser todas críticas. Sirven para mapear el entorno, anticipar cambios y actualizar la TdC cuando se mueven.

- Dependencias Requisitos con terceros o con programas/recursos ajenos sin los cuales no se materializan productos o resultados. Ejemplos de ello podrían ser dependencia de cofinanciamiento, acceso a equipamiento o data.
- Riesgos: Eventos o condiciones (internas o externas) que pueden impedir o retrasar los vínculos causales, incluidos los derivados del incumplimiento de supuestos o dependencias. Se asocian a medidas de mitigación y a puntos de control en el plan de monitoreo y evaluación.



DATOS EN SISTEMAS DE MONITOREO Y EVALUACIÓN (M&E)

En una evaluación sistémica, los datos dejan de ser solo insumos técnicos y se convierten en un recurso estratégico. Su gestión adecuada permite rendir cuentas, aprender colectivamente y orientar decisiones con evidencia, en un sistema de Monitoreo y Evaluación (M&E) que incluye interacciones y contribuciones dentro del ecosistema de CTCl.

2.1 Datos e indicadores para la evaluación sistémica

Medir CTCl con enfoque sistémico significa observar trayectorias y contribuciones, no solo resultados aislados. Los datos permiten seguir tres planos que se entrelazan:

- Outputs: señales tempranas que muestran activación y dirección (proyectos, prototipos, publicaciones, servicios);
- Outcomes entendidos como reconfiguraciones que sostienen cambios en el tiempo, tales como nuevas capacidades, adopción tecnológica, vinculación ciencia-empresa, encadenamientos; e
- Impactos de mayor escala y plazo (productividad, sostenibilidad, salud, inserción en cadenas globales); siempre entendidos como contribución colectiva del portafolio y del ecosistema.

Para leer estas trayectorias es necesario cruzar distintos tipos de evidencia. Los datos cuantitativos aportan magnitud y ritmo, los cualitativos explican mecanismos y contextos, y los contextuales sitúan

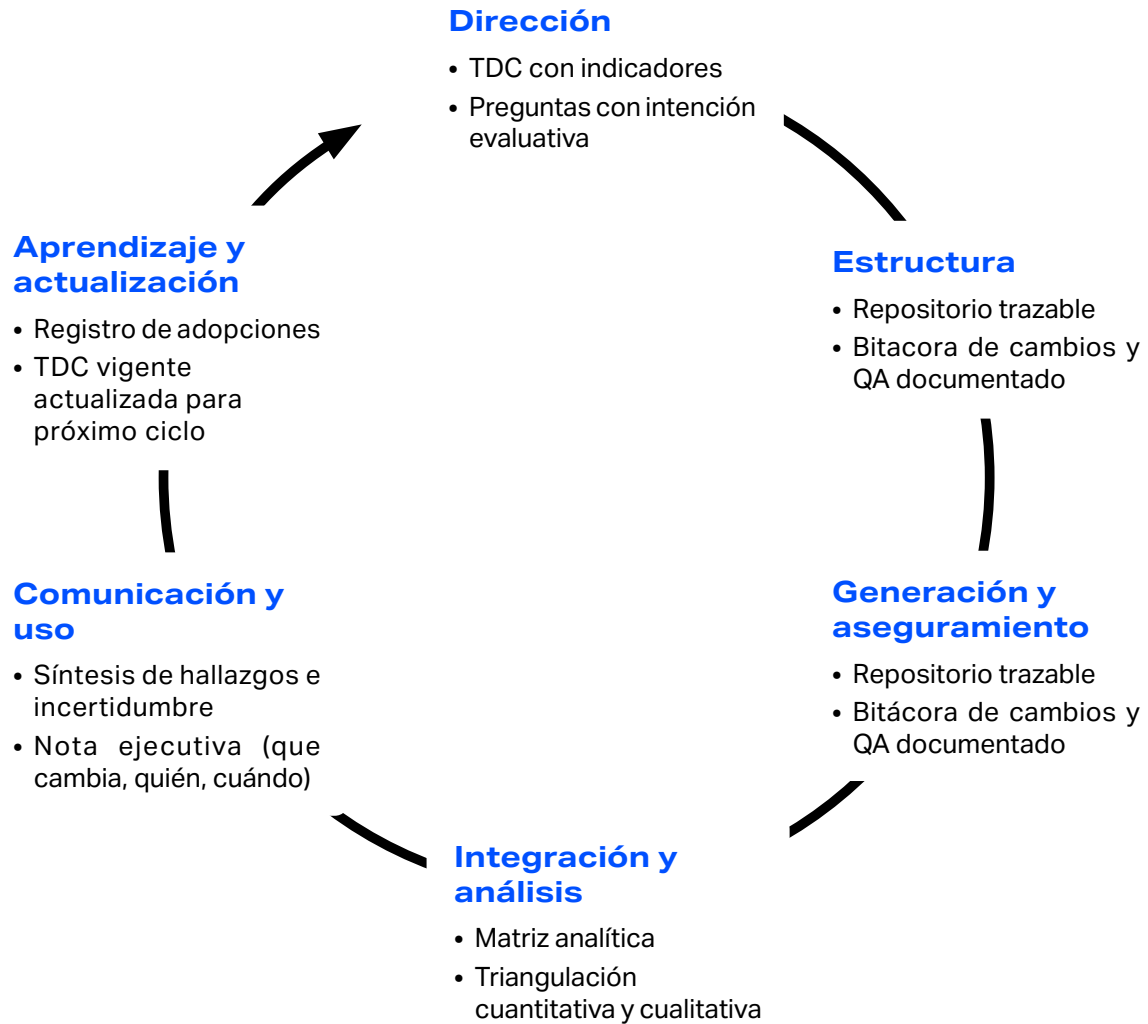
las condiciones de posibilidad (regulación, shocks, territorios, mercados). Al mismo tiempo, se observan funciones del sistema (legitimación, formación de mercados, intermediación, circulación de talento) y activos intangibles (absorción, reputación, densidad relacional, confianza), que se manifiestan en trazas como coautorías, movilidad o contratos. Son estos factores los que explican por qué algunos entornos convierten con mayor rapidez los outputs iniciales en outcomes e impactos duraderos.

2.2 Ciclo de datos integrado: monitoreo y evaluación

La gestión de datos en CTCl es una función estratégica: orienta decisiones, permite aprender y sostiene la direccionalidad de portafolios y misiones. No son dos circuitos aislados, sino un mismo sistema con dos ritmos: monitoreo (pulso continuo) y evaluación (miradas profundas).

El sistema de datos en CTCl no opera como circuito aislado, sino como infraestructura compartida que alimenta ambos ciclos. En la práctica, el monitoreo lo conduce la agencia ejecutora (por ejemplo, ANID), coordinada con el ministerio que fija prioridades y ventanas de decisión; administra el repositorio, los metadatos y el aseguramiento de calidad. Las evaluaciones pueden ser impulsadas por el ministerio/agencia para gestión y aprendizaje, o coordinadas por DIPRES cuando se vinculan al ciclo presupuestario, aplicando sus estándares de independencia y trazabilidad.

El ciclo, representado en seis etapas, muestra cómo los datos transitan desde la definición de preguntas hasta la institucionalización de aprendizajes, articulando un flujo que combina rutinas de gestión y análisis más profundos:

DIAGRAMA 3: ETAPAS DEL CICLO DEL DATO INTEGRADO

FUENTE: ELABORACIÓN PROPIA

1. En Dirección, los indicadores definidos para el monitoreo derivan de preguntas evaluativas (se plasma en la TDC), lo que asegura que cada dato tenga una intención de uso y alimente hipótesis de cambio.
2. Estructura configura la arquitectura común de datos, con acuerdos de transferencia e identificadores únicos, permitiendo interoperabilidad siempre con resguardo ético.
3. En Generación y aseguramiento se consolida la trazabilidad: huella que habilita auditoría y reutilización. Series rutinarias y repositorios auditables con bitácora y controles documentados deja evidencia de quién, cuándo y cómo se produjo y depuró cada registro.
4. La Integración y análisis conecta la operación con la interpretación a través de mecanismos para que la evidencia generada sea verificable y explicativa. Los tableros del seguimiento se alimentan de una matriz analítica que encadena: pregunta–TDC–dato–método–producto, y la evaluación añade enfoques mixtos y lectura multinivel cuando corresponde estimar efectos.
5. En la Comunicación y uso se convierte la evidencia en insumo de decisión (diseñada conforme al perfil o audiencia), con mensajes claros y oportunos. Estrategias: informes breves y paneles interactivos para el monitoreo.
6. Finalmente, en Aprendizaje y actualización, se registra lo adoptado y se actualiza la TDC, asegurando memoria institucional y un inicio más avanzado para el siguiente ciclo.

De este modo, el sistema de datos se convierte en el puente metodológico entre el monitoreo y la evaluación: provee consistencia, trazabilidad y capacidad de uso en ambos niveles, evitando duplicidades y favoreciendo que cada indicador y cada análisis estén alineados con las decisiones estratégicas del sistema CTCl.



2.3 Principios para la gestión de data en CTCI

Los principios de datos para el seguimiento y evaluación orientan el uso de la evidencia en función de su propósito. No se trata solo de acumular información, sino de establecer criterios comunes que den coherencia y sentido a lo que se mide y cómo se utiliza.

Relevancia

Cada indicador debe nacer de una pregunta de decisión y anclarse a la TDC: esto significa que medimos lo que efectivamente orienta diseño, ajuste o cierre de programas. En otras palabras, los datos recopilados deben ser pertinentes al contexto y objetivos del programa, evitando métricas irrelevantes o de "relleno". La buena práctica internacional subraya que los indicadores han de reflejar la consecuencia de la actividad y alinearse con la cadena causal; de lo contrario, es ruido. En la misma línea, el estándar chileno recuerda que los "datos aislados" valen poco si no se conectan con decisiones: un indicador relevante traza, de la métrica a la pregunta, y de la pregunta a la decisión, justificando el esfuerzo de medición por su valor de uso

Calidad

Combina precisión, completitud, oportunidad y coherencia a lo largo del tiempo. En la práctica: registros sin errores ni vacíos, series con cambios metodológicos documentados (metadatos) y controles de consistencia temporal para evitar lecturas engañosas. Marcos nacionales la definen con esos cuatro atributos mínimos y advierten sobre los sesgos de duplicados, faltantes o atípicos; además promueven QA independiente y, cuando proceda, revisión por pares, para reforzar credibilidad.

Oportunidad

Implica que los datos estén disponibles en el momento en que se toman las decisiones. No basta con contar con información precisa: si llega con rezago, pierde valor. Marcos como el usado por el Department for Business, Energy and Industrial Strategy de UK (BEIS) y DIPRES destacan que la utilidad de un indicador depende de su puntualidad y periodicidad, ajustadas a los hitos de política (diseño, reformulación, ciclos presupuestarios). Cumplir calendarios de actualización, es señal de un sistema de M&E capaz de proveer evidencia vigente para planificar, ajustar programas y responder con agilidad.

Accesibilidad y usabilidad

Evidencia abierta y comprensible: publicación de bases, metodologías y glosarios; visualizaciones y diccionarios que permitan interpretar; y protocolos de difusión para que los hallazgos lleguen al usuario indicado en el momento cuando debe tomar decisiones. Experiencias comparadas muestran el valor de publicar evaluaciones y habilitar acceso responsable, fortaleciendo transparencia y aprendizaje.

Comparabilidad

Los datos deben ser comparables en territorios, sectores y tiempo. Esto exige definiciones y métodos estandarizados, series rastreables y diccionarios compartidos que prevengan comparaciones no equivalentes e inconsistencias de medición. La coherencia interna entre indicadores y la alineación con marcos comunes permiten agregar, contrastar y seguir tendencias con confianza.

Proporcionalidad y legitimidad

El esfuerzo de M&E debe ser acorde a la escala y riesgo de cada intervención: ni sobrecargar con métricas inútiles ni subestimar programas críticos. La legitimidad exige resguardo ético y de privacidad, minimizar cargas a informantes y, cuando corresponda, revisión externa/participación. Un plan de datos por intervención con evaluación costo-utilidad ayuda a equilibrar información necesaria y cuidado de las personas.

Estos principios deben aplicarse con sentido práctico. Las instituciones operan con capacidades desiguales, sistemas heredados y restricciones legales que hacen inviable una estandarización inmediata. Avanzar hacia una gestión de datos coherente en CTCl requiere pragmatismo: priorizar lo esencial, construir sobre lo que ya existe y fortalecer gradualmente las condiciones técnicas y de gobernanza. Más que buscar el modelo perfecto, se trata de asegurar trazabilidad y valor de uso en cada paso, consolidando aprendizajes que permitan escalar sin perder la conexión con la realidad institucional.



METODOLOGÍAS DE EVALUACIÓN

No hay un método de evaluación más efectivo que otros en todos los escenarios. La elección del método de evaluación será dependiente del tipo de intervención, de las preguntas de evaluación, de las prioridades de quienes toman las decisiones o de los actores involucrados, de los datos disponibles o potenciales a recabar y del tiempo y recursos con los que se cuente. Cada método tiene sus beneficios y sus limitaciones.

En el contexto de esta Guía se consideran las siguientes definiciones para referirse a: aproxima-

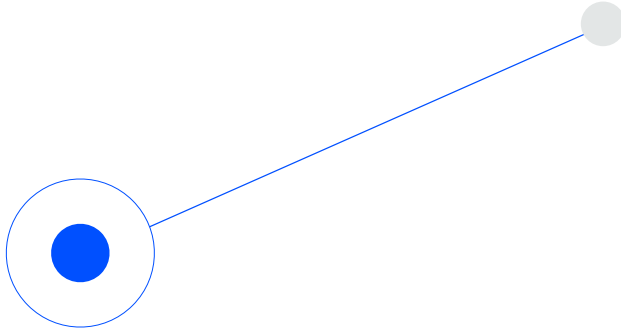
ción, método de evaluación y técnica de evaluación o investigación (Befani, 2020). La aproximación se refiere a la concepción teórica general relativa a los métodos de investigación o evaluación, de su sentido ontológico. Si bien estas concepciones son subyacentes a cada método, no se profundizará en ellas para los usos de este documento, sino que se considerarán los métodos y técnicas en su sentido práctico. En el caso de los métodos, se refiere a los procesos con los que se dará respuesta a las preguntas de evaluación. La agrupación y el listado de los métodos será el foco de esta sección. Finalmente, las técnicas de investigación o evaluación son los procedimientos para levantar o producir la información y los datos que serán utilizados en los métodos de evaluación, estos son generales a la investigación social y son, principalmente, los siguientes: Encuestas, entrevistas, grupos focales, datos de monitoreo, técnicas de observación, técnicas deliberativas o participativas y estudios de caso.

Para realizar una revisión de los métodos de evaluación, se seguirá la siguiente agrupación: Métodos Experimentales y Cuasiexperimentales, Métodos Basados en Teoría, Métodos Costo-Beneficio y Métodos de Síntesis de evidencia o Meta-Análisis.

3.1 Experimental y cuasi-experimental

Son útiles para generar una estimación de impacto, es decir, el tamaño y la significancia del cambio producido por una intervención. El principio central es el uso de un grupo "contrafactual" para comparar con el grupo intervenido. Por sí solo no establece cómo y por qué ocurrieron los impactos. Tiene un costo mayor² y requerimientos específicos de datos.

² Si es que es necesario levantar información primaria mediante técnicas de investigación.



Experimento aleatorio Controlado (Random Controlled Trials)

Se genera un grupo a intervenir y un grupo de control, los individuos para cada grupo se eligen de forma aleatoria dentro de la población objetivo. Se controlan al inicio y al final de la intervención. De esta manera, se pueden comparar los resultados tanto del grupo que recibió la intervención como del que no la recibió y así determinar los efectos netos. Sus aspectos positivos es que es grupo contrafactual menos sesgado, debido al procedimiento aleatorio. En sus aspectos negativos es que es poco factible en muchos casos reales y tiene implicancias éticas en determinados contextos.

Para profundizar:

- BetterEvaluation.org (2019). Randomised Controlled Trial. Available at: www.betterevaluation.org/en/plan/approach/rct
- Duflo E., R. Glennerster y M. Kremer (2008). "Using Randomization in Development Economics Research: A Toolkit", en T. P. Schultz y J. A. Strauss (Eds.). Handbook of Development Economics, Vol. 4, Elsevier.
- BID (2016). La evaluación de impacto en la práctica.

Ejemplo de aplicación

- Evaluación de impacto de Chile va!, encuentro de jóvenes con la ciencia y la tecnología (2014). Ver aquí

Regresión Discontinua

En intervenciones que tienen una variable que puntúa o genera un ranking de los sujetos a intervenir, es decir, un puntaje de selección, se puede asumir que quienes, dentro de un margen en torno al puntaje de corte, cumplen con los requisitos y quienes no los cumplen, son grupos similares y comparables. Así se genera el grupo contrafactual o de control con quienes quedan por debajo del puntaje de corte. El análisis luego resulta similar al anterior, comparando los grupos en las variables de interés antes y después de la intervención. Sus aspectos positivos son que el grupo contrafactual se asume cerca de la distribución aleatoria y que se genera "naturalmente" por la naturaleza del programa al asignar puntajes a los individuos y una línea de corte. En sus aspectos negativos es que no se puede generalizar las conclusiones ni al grupo con puntajes más altos del margen definido ni al grupo con menor puntaje.

Para profundizar:

- Thistlethwaite, D.; Campbell, D. (1960). "Regression-Discontinuity Analysis: An alternative to the ex post facto experiment", *Journal of Educational Psychology*, 51 (6) pp.309-317. doi:10.1037/h0044319.
- Lee, D. y T. Lemieux (2010). "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

Ejemplo de aplicación

- Evaluación de impacto de la inversión de Corfo en instrumentos de fomento de la innovación. (2022) Ver aquí
- Evaluación de Resultados e Impacto FONDECYT (2021) Ver aquí



Diferencia en diferencias

El grupo contrafactual se construye al hacer seguimiento en función de las variables de interés, las cuales van a ser afectadas por la intervención, tanto del grupo a intervenir, como de otro que se comporte de la misma forma antes de la intervención. Así, según el comportamiento anterior de ambos grupos se genera una tendencia, que deberían seguir ambos grupos. Posterior a la intervención se miden las mismas variables en ambos grupos, confirmando la tendencia en el grupo contrafactual y observando la diferencia que existe con la tendencia del grupo que recibió la intervención, pudiendo calcular así el efecto neto. Sus aspectos positivos es que es una buena alternativa a otros métodos de generar un grupo contrafactual. Entre sus aspectos negativos, está la necesidad de tener una buena cantidad y calidad de datos de los grupos a analizar, tanto antes como después de la intervención.

Para profundizar:

- Gertler, Martinez, Premand, Rawlings, & Vermeerch, 2011
- Gruber, J. (1994). "The Incidence of Mandated Maternity Benefits". *American Economic Review* 84(3): 622-641.

Ejemplo de aplicación

- INFORME DE EVALUACIÓN PROGRAMAS DE BECAS DE POSGRADO (2019). ANII Ver aquí

Emparejamiento de Puntajes de Propensión (Propensity Score Matching)

Se construye un contrafactual en base al puntaje de propensión, que es el puntaje que estima la probabilidad de ser intervenido, el cual va de 0 a 1, donde 1 es 100% de probabilidad de ser intervenido. Tanto el grupo intervenido como el contrafactual deben tener individuos con puntajes de propensión similares, así, se generan ambos grupos con condiciones similares en los factores relevantes para la intervención. En sus aspectos positivos destaca que puede estimar el impacto promedio de la intervención en toda la población beneficiaria y no de un subconjunto de la misma. Sus aspectos negativos son que necesita datos en considerable cantidad y calidad, y que sean recabados antes de la intervención o que no sean dependientes del tiempo.

Para profundizar:

- Ipsos Mori Social research institute (2018) Sheffield Hallam University and London Economics.
- University of North Carolina, University of Ghana, Food & Agricultural Organisation of the United Nations (2014) Livelihood Empowerment Against Poverty Program Impact Evaluation, North Carolina.
- Caliendo, M. y S. Kopeinig (2008) "Some Practical Guidance for the Implementation of Propensity Score Matching". *Journal of Economic Surveys* 22(1): 31-72.

Ejemplo de aplicación

- Evaluación de resultados ex post del programa de difusión tecnológica (PDT) (2018). Ver aquí
- Evaluación de Impacto Centros Basales (2021) Ver aquí. A nivel de investigador se utilizó el método de emparejamiento de puntajes de propensión.
- Diseño de una metodología con el objeto de medir los cambios en la cultura de la ciencia, tecnología e innovación en niños, niñas y jóvenes, en el marco del proyecto piloto de educación y cultura cti, informe final. (2016). Ver aquí
- Estudio crédito tributario para la I+D (2013). Ver aquí
- Public support to firm-level, innovation: an evaluation of the fontec program (2007) Ver aquí



Método de Control Sintético

El contrafactual es un clon sintético utilizando una gran cantidad de datos secundarios históricos. Para hacerlo, se generan promedios ponderados de las variables de interés relacionadas con las características de la población intervenida y de sus resultados. La comparación entre la población intervenida y el clon sintético permite estimar el impacto de la intervención. Se usa principalmente cuando son pocos los casos intervenidos o tratados por el programa. Su limitación estriba en que el método sólo es posible cuando se puede establecer una relación duradera o histórica entre el comportamiento del grupo tratado con los casos que se usan para generar el grupo de control.

Para profundizar:

- Abadie, A.; Diamond, A.; and Hainmueller, J (2012) Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program, *Journal of the American Statistical Association*, 105: 490, pp. 493-505.
- Craig, P. (2015) Synthetic controls: a new approach to evaluating interventions, *What Works Scotland*, Working paper.

Ejemplo de aplicación

- Estudio de evaluación de la ley de incentivo tributario a la inversión en i+d (2017) Ver aquí
- Evaluación de Impacto Centros Basales (2021). A nivel de instrumento se aplicó el Método de control sintético.

VARIABLES INSTRUMENTALES

El contrafactual se produce al usar una variable que permita predecir la probabilidad de ser intervenido pero que no tenga impacto en los resultados de la intervención. En sus aspectos positivos es que es un método posible cuando otros diseños experimentales o cuasi experimentales no se pueden llevar a cabo. En sus aspectos negativos se encuentra la dificultad de encontrar la variable instrumental, tanto por el requerimiento de que no influya en los resultados de la intervención y porque es difícil definirla antes de la intervención porque no se puede asegurar totalmente que no tendrá impacto en sus resultados. Otro aspecto negativo es que la estimación del impacto es "local" ya que solo se aplica para el grupo que fue intervenido y los individuos que se encuentran muy cerca de ese grupo.

Para profundizar:

- Maurin, E. and McNally, S. (2008). Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students. *Journal of Labor Economics* 26,1: pp.1-33.

Ejemplo de aplicación

- The macroeconomic effects of the fraunhofer-gesellschaft (2021). Ver aquí

3.2 Métodos Basados en Teoría

Pueden utilizarse para la evaluación de impacto respondiendo si la intervención causó un impacto, cómo y por qué ocurrió, cómo el contexto pudo haber influido, aunque no dan estimaciones precisas de la magnitud del efecto. Permite una mirada amplia y tiene un costo menor.

Un ejemplo de un método que se acerca a los basados en teoría, se usó para el estudio: Evaluación de impacto de proyectos y giras de innovación de FIA (2015). Ver aquí

Análisis de contribución

Se trata del análisis en detalle, y entregando evidencias, de la teoría del cambio del programa y sus relaciones causales establecidas. Además, hace una revisión de los supuestos y otros factores que podrían estar influyendo. Esta metodología defiende que si se pueden confirmar las relaciones causales de la teoría del cambio con evidencia, es razonable atribuir los resultados a la intervención. Entre sus aspectos positivos se encuentran el poder confirmar y actualizar la teoría del cambio y que es una buena opción cuando no se desea o no es posible realizar métodos experimentales. Entre sus limitaciones está el no entregar evidencias definitivas de las relaciones causales. Otra limitación es que trabaja sobre efectos promedio, no se recomienda su uso cuando existen distintos resultados para distintos subgrupos.

Para profundizar:

- Mayne, J. (2008) Contribution Analysis: An approach to exploring cause and effect. ILAC methodological brief.
- Remnant, F.; Copestake, J. (2015) The Qualitative Impact Assessment Protocol (QUIP). Centre for Development Studies, University of Bath, Bath, UK pp. 1-9.

Ejemplo de aplicación

- Support to large enterprises Final Report- Work Package 4 (2016). Ex post evaluation of Cohesion Policy programme 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF). Ver aquí

Rastreo de procesos

Metodología que se aplica en un solo caso, en donde se generan hipótesis causales que podrían explicar los resultados de la intervención, y estas se contrastan, buscando evidencia tanto de si la hipótesis fuese verdadera como si ésta fuese falsa. Entre sus aspectos positivos están que no se necesita contrafactual y que puede ser realizado posterior a la realización de la intervención. Entre sus aspectos negativos está que el centrarse en una hipótesis causal, tanto al verificar como al refutar, no significa que no pueda haber otra, por lo que las explicaciones alternativas deben ser cuidadosamente consideradas.

Para profundizar:

- Befani, B. & Stedman-Bryce, G. (2017) Process Tracing and Bayesian Updating for impact evaluation. Evaluation, 23(1), pp.42-60.

Ejemplo de aplicación

- Mid-term Evaluation and Learning Exercise of the Chile Self-Supply Renewable Energy Project (2023) Ver aquí

Actualización bayesiana (Bayesian updating)

Es un procedimiento para formalizar y reforzar las afirmaciones hechas en otras metodologías como rastreo de procesos y análisis de contribución. Utiliza el teorema de Bayes para calcular las probabilidades de que una afirmación de contribución sea verdadera. Se deben generar estimaciones de en qué medida las distintas evidencias disponibles aumentan la probabilidad de que las atribuciones causales sean verdaderas o falsas, generando una transparencia formalizada de los criterios de los evaluadores. Entre sus aspectos positivos están la capacidad de reforzar las afirmaciones de otras metodologías basadas en teoría, poder manejar un rango amplio de evidencias de distinta calidad. Entre sus aspectos negativos está el que el uso de probabilidades puede hacer poco claras las conclusiones para los distintos tipos de usuarios.

Para profundizar:

- CECAN (2016) Testing Contribution Claims with Bayesian Updating. A CECAN Evaluation and Policy Practice Note for policy analysts and evaluators Note No. 2.1. Winter 2016

Ejemplo de aplicación

- Bayesian Impact Evaluation with Informative Priors. An Application to a Colombian Management and Export Improvement Program. (2023) Ver aquí
-

Evaluación realista (Realist evaluation)

Se trata de generar hipótesis causales de cómo los individuos aprovechan o no las intervenciones o programas, por lo que profundizan incluso en procesos psicológicos. Estas hipótesis causales son llamadas "mecanismos", que pueden o no funcionar también muy dependientes del contexto. Esta metodología trata de entender qué es lo que funciona, para quien o quienes y en qué circunstancias. Luego de definido el "mecanismo" busca las evidencias que lo pueden respaldar o refutar. Entre sus aspectos positivos está la profundización que busca en la forma en la que los individuos utilizan o no las intervenciones o programas. Entre sus limitaciones está el no generar estimaciones cuantitativas de los efectos de las intervenciones y el requerir bastante tiempo y recursos para su aplicación.

Para profundizar:

- Gill Westhorp (2014) Realist Evaluation: An Introduction. Methods Lab Overseas Development Institute London (Westhorp, 2014)

Ejemplo de aplicación

- Evaluation of the transitional arrangements for demand-side response (2022). Ver aquí
-

Análisis comparativo cualitativo

Se trata de una comparación cualitativa entre múltiples casos para entender los patrones o regularidades que se presentan en los resultados de las intervenciones. Permite identificar una gama de distintos factores que se relacionan con el éxito o fracaso de las intervenciones. Entre sus aspectos positivos se encuentra que permite entender por qué una intervención funcionó en un contexto pero en otro no. Entre sus limitaciones se encuentra que requiere una cantidad amplia de casos para poder aplicarse.

Para profundizar:

- Blackman, T., Wistow, J. & Byrne, D. (2011) A Qualitative Comparative Analysis of factors associated with trends in narrowing health inequalities in England. *Social Science & Medicine*, 72, pp. 1965-1974.
- Welle, K., Williams, J., Pearce J. & Befani, B. (2015) Testing the Waters: A Qualitative Comparative Analysis of the Factors Affecting Success in Rendering Water Services Sustainable Based on ICT Reporting. Wateraid, itad & IRC.

Ejemplo de aplicación

- Evaluation of the Energy Entrepreneurs Fund. Final Report (2023). Ver aquí

3.3 Costo-beneficio

Comparan los beneficios y los costos de las intervenciones, incluyendo efectos adversos o no intencionados. No siempre es evidente qué es un costo y qué un beneficio.

Análisis de Costo-Efectividad social

Analiza y cuantifica los costos de intervenciones que cuentan con las mismas métricas de sus resultados, con el objetivo de seleccionar la mejor opción. Se utiliza cuando no es posible o deseable cuantificar los beneficios de una intervención.

Para profundizar:

- Kumaranayake, L.; Vickerman, P.; Watts, C.; Guinness, L. (2002) Cost-effectiveness analysis: aiding decision-making in HIV prevention in the Ukraine. Report prepared for British Council/DFID.
- Lagarde, M.; Blaauw, D.; Cairns, J. (2012) Cost-effectiveness analysis of human resources policy interventions to address the shortage of nurses in rural South Africa. *Social Science and Medicine* 75 (5) pp. 801-806. [DOI: 10.1016/j.socscimed.2012.05.005]

Ejemplo de aplicación

- Servicio de consultoría para la elaboración del estudio de productividad científica de los centros del programa iniciativa científica milenio (2016). Ver aquí

Análisis de Costo-Beneficio social

Analiza y cuantifica tanto los costos como los beneficios de las intervenciones, lo que permite comparar distintas intervenciones con distintas métricas de sus resultados. Entre sus limitaciones está la necesidad de contar con datos en cantidad y calidad.

Para profundizar:

- HM Treasury (2018) The Green Book: Central Government Guidance on Appraisal and Evaluation. London. Crown Copyright. Available at: www.assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685903/The_Green_Book.pdf

Ejemplo de aplicación

- Análisis Costo Beneficio (ACB) Programa Centros de Excelencia Internacional (CEI) (2019). Ver aquí

3.4 Meta-análisis: Síntesis de evaluaciones pasadas

Revisión sistemática de evidencia evaluativa acumulada, utilizando algún criterio de selección tipo de programa, resultados o enfoques, etcétera. Utilizadas para extraer lecciones y orientar decisiones futuras. Es menos costoso, aunque depende de la disponibilidad de evaluaciones y datos.

Revisión sistemática

Se trata de una revisión exhaustiva de los estudios referidos al tema o pregunta de investigación, que permite la revisión y consolidación de la evidencia que surge de muchos de ellos. Tiene por objetivo, identificar todos los estudios relevantes del área, para lograr una revisión no sesgada, transparente y replicable. Es necesario analizar la calidad de las evidencias de los estudios. Finalmente, se interpretan los hallazgos.

Entre sus limitaciones está que una revisión exhaustiva de un área requiere bastantes recursos.

Para profundizar:

- Khan, K.; Kunz, R.; Kleijnen, J. and Antes, G. (2003). Five steps to conducting a literature review. *Journal of the Royal Society of Medicine*, 96 (3) pp. 118–121.
- Barlow, J; Fisher, J. and Jones, D. (2012). Systematic review of models of analysing significant harm. Department for Education, London.

Metaanálisis

Genera una integración estadística de los resultados cuantitativos de distintos estudios, teniendo como resultado, por ejemplo, un promedio del efecto de las intervención o expresado como rango de efecto. La calidad del Meta-análisis dependerá de la calidad de los estudios usados como fuente de información.

Para profundizar:

- Glass, G. (1976). Primary, Secondary and Meta-Analysis of Research. *Educational Researcher*, 5, (10) pp. 3-8
- Colliver, J.; Kucera, K. and Verhulst, S. (2008). Meta-analysis of quasi-experimental research: are systematic narrative reviews indicated?. *Medical Education*, 42 (9) pp. 858-865

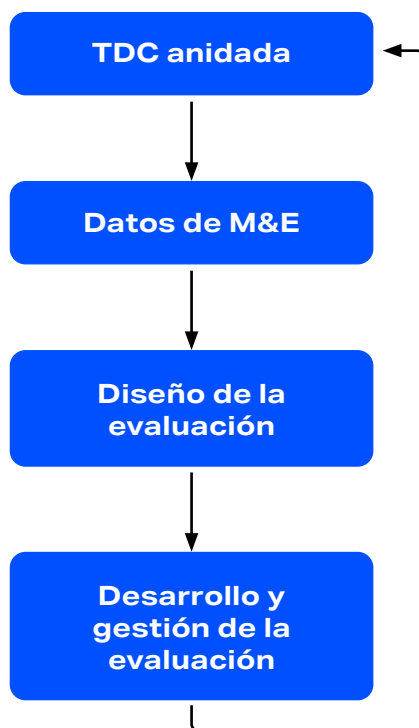
Ejemplo de aplicación

- The Impact and Effectiveness of Innovation Policy: Evidence from Middle-Income Countries (2024). Ver aquí

ELEMENTOS CLAVE PARA UNA EVALUACIÓN CON ENFOQUE SISTÉMICO EFECTIVA

Este apartado organiza los elementos esenciales como una lista de cotejo para orientar, paso a paso, una evaluación con enfoque sistémico en CTCL. Cada dimensión se presenta en el orden lógico del proceso, con su definición, orientaciones prácticas y alternativas o plan B frente a obstáculos.

En un escenario ideal, la Teoría de Cambio define los objetivos, supuestos e interacciones; la data de monitoreo y evaluación se estructura en función de esa lógica; luego se formula el diseño de la evaluación con preguntas y métodos claros; y finalmente, la gestión y desarrollo aseguran la implementación y uso de resultados.



En la práctica, muchas evaluaciones comienzan desde el diseño sin contar con TDC o con datos suficientes. Por eso explicitamos este orden como referencia: el estándar hacia el que conviene avanzar para que el diseño y la implementación sean más claros, ordenados y efectivos.

Si falta TDC o los datos son insuficientes, no se detiene el proceso: arma un mínimo viable (TDC breve, inventario de datos con brechas y supuestos explícitos) y avanza con un plan concreto para cerrar esas brechas mientras se ejecuta la evaluación. Cada etapa puede devolver a la anterior cuando surgen hallazgos o restricciones relevantes.

Considerando lo anterior, es que previo a la evaluación, se invita a hacer un encuadre: una pausa para ponernos de acuerdo. ¿Para qué vamos a evaluar? ¿Qué decisiones vienen y cuándo? ¿Quiénes necesitan los resultados y cómo se involucrarán? ¿Con qué márgenes contamos en términos de tiempo, datos y normas? Dejamos estas respuestas por escrito, así este acuerdo inicial guía la TDC y ordena el resto del proceso.

En las páginas siguientes, la lista de cotejo se organiza en dimensiones con respectivos elementos, siguiendo esta cadena. En cada punto encontrarás qué verificar, cómo aplicarlo y opciones de contingencia ante limitaciones de información, tiempo o gobernanza. El objetivo es conectar datos, análisis y decisiones de forma trazable y oportuna para sus distintos usuarios.

Elementos para una evaluación con enfoque sistémico en CTCl

Teoría de Cambio anidada Mapa causal que describe cómo una intervención específica aporta al cambio esperado de un programa, política o sistema mayor, alineando sus rutas causales y supuestos con las escalas superiores.

1. Cómo construirla: 6 pasos principales, pasando desde el diagnóstico hasta prueba de consistencia
2. Anidamiento y articulación sistémica en multinivel
3. Temporalidad y plan de evaluación
4. Participación y validación
5. Actualización y mínimo viable

Data de monitoreo y evaluación Sistema de datos capaz de traducir la TDC en indicadores, gestiona y gobierna datos y asegura su uso en decisiones.

1. Planificar desde la Teoría de Cambio
2. Captura y gobernanza de datos
3. Analizar e interpretar: de información a evidencia
4. Comunicar, decidir y mejorar: cerrar el ciclo

Diseño de la evaluación sistémica Aspectos clave para la definición de objetivos, preguntas, métodos y roles, con proporcionalidad y rigurosidad metodológica.

1. Formulación de objetivos y preguntas de evaluación
2. Requerimientos de información de hallazgos/resultados para la toma de decisión
3. Enfoque y método
4. Revisión de pares y expertos independientes

Desarrollo de la evaluación sistémica Descripción de la gestión y aspectos de calidad a los que atender en el "paso a paso" de la ejecución de la evaluación.

1. Gobernanza y resguardos
2. Planificación
3. Diseño proporcional
4. Consolidación y estandarización de datos
5. Trabajo de campo: ejecución conforme a diseño metodológico
6. Análisis y síntesis
7. Uso de resultados: recomendaciones y difusión



TEORÍA DE CAMBIO ANIDADA

La TDC es el mapa causal compartido que orienta por qué intervenir, cómo se espera que ocurra el cambio y bajo qué condiciones eso resulta plausible. Sirve para tomar decisiones: acota el alcance, prioriza preguntas e indicadores. Al anidarse en estrategias, misiones y portafolios, muestra la contribución del programa dentro del sistema y explícita dependencias con otros instrumentos.

1.1 Cómo construirla: 6 pasos principales

1) Diagnóstico situado

Delimita el problema u oportunidad, su magnitud y evolución; identifica causas inmediatas y subyacentes y el costo de no actuar. Describe el contexto regulatorio y de mercado, capacidades institucionales, cultura y territorio. Sin este anclaje, la causalidad queda sin sustento o "en el aire".

2) Propósito e impacto

Precisa el cambio buscado y el impacto al que se aspira contribuir en una población determinada. Define el nivel de trabajo (instrumento, programa, portafolio o misión) y cómo se anida en marcos superiores. La TDC se basa en contribución, no en atribución exclusiva.

3) Rutas causales y cadena de resultados

Representa esta cadena en un diagrama legible con una narrativa breve, incluyendo rutas paralelas y retroalimentaciones posibles.

- Secuencia lógica: insumos y actividades → productos → resultados intermedios → impacto.
- En cada enlace aplica: "si A ocurre, debiera observarse B, porque opera el mecanismo C y siempre que se cumplan las condiciones C".

4) Condiciones de cambio

Por cada enlace A→B, explicita supuestos, riesgos, factores externos y dependencias. Define señales tempranas por ruta y asigna responsables de su seguimiento y de la activación de rutas.

5) Indicadores y preguntas evaluativas

Para cada resultado priorizado, define al menos un indicador y una pregunta. Completa una matriz con: resultado, pregunta, indicador, fuente, frecuencia, responsable, hito de decisión, limitaciones.

6) Prueba de consistencia

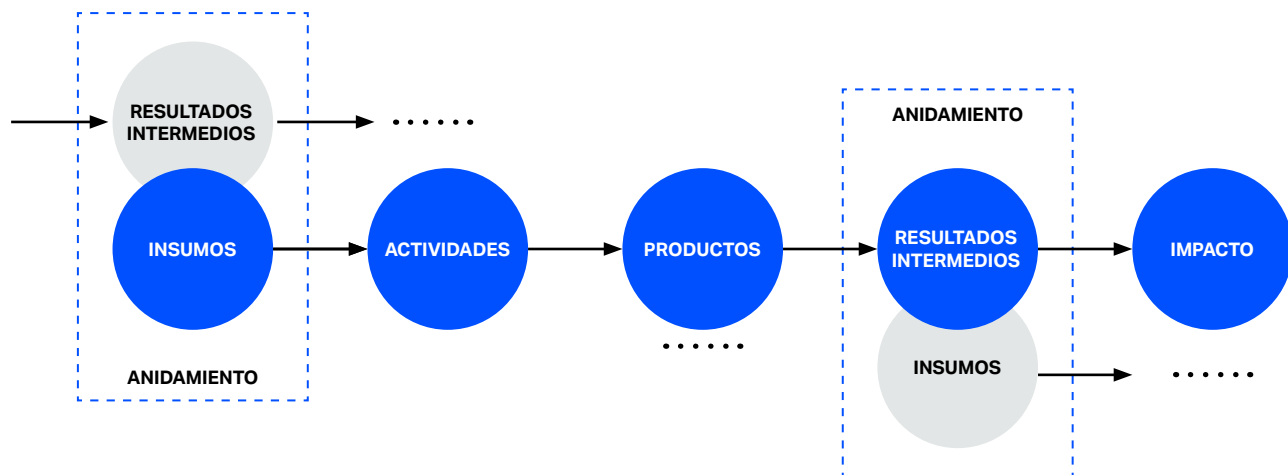
Aplica teoría negativa (por qué podría no ocurrir cada salto), contrasta con evidencia previa (literatura, evaluaciones, series) y valida con implementadores, beneficiarios y pares. Documenta grados de confianza, si es alto, medio o bajo, y sus razones (calidad de datos, consistencia entre fuentes, sensibilidad a supuestos). Ajusta la TDC y registra la versión.

1.2 Anidamiento y articulación sistémica en multinivel

Una evaluación con enfoque sistémico requiere comprender cómo cada instrumento o proyecto se inserta en una cadena causal más amplia—ya sea un programa, portafolio o misión—, y cómo sus resultados se relacionan con los de otros instrumentos que operan en el mismo espacio. Esta lectura multinivel permite mantener coherencia entre objetivos,

identificar dependencias y promover decisiones coordinadas. En este marco, se habla de Teoría de Cambio anidada cuando las cadenas causales de diferentes niveles se ordenan y conectan de manera explícita, y de articulación cuando existen vínculos operativos entre ellas (Magenta, BEIS).

DIAGRAMA 4: ANIDAMIENTO EN LA TEORÍA DE CAMBIO



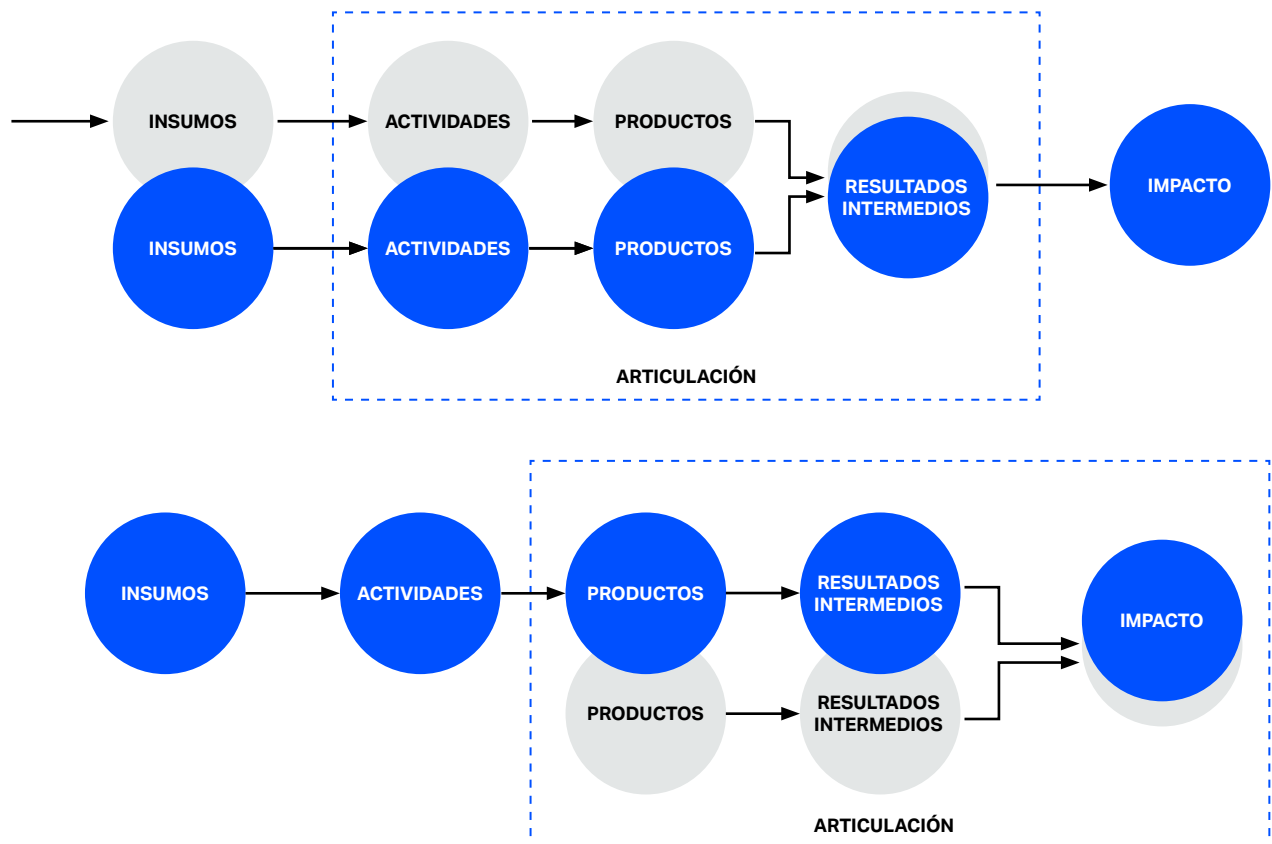
Una Teoría de Cambio anidada ubica la cadena causal de un instrumento o proyecto dentro de la cadena de su programa o portafolio, y a su vez dentro de la misión o estrategia superior. El anidamiento ocurre tanto al inicio de la cadena, al definir propósitos, supuestos y condiciones que orientan la acción, como al final, cuando los resultados e impactos del nivel inferior se acumulan como logros verificables a nivel de programa o misión (HM Treasury, 2020).

De este modo, los productos y resultados intermedios de cada instrumento no se pierden ni se duplican, sino que se encadenan de manera trazable con los objetivos de niveles superiores, favoreciendo la coherencia entre diseño, seguimiento y evaluación (BEIS, 2020).

La articulación refiere a las relaciones horizontales que se producen entre distintos instrumentos o proyectos dentro de un mismo portafolio o misión. A diferencia del anidamiento, que estructura las relaciones verticales, la articulación ocurre en los eslabones intermedios de la cadena, donde los productos o resultados de un instrumento habilitan o fortalecen los de otro. También puede existir convergencia, cuando varias cadenas contribuyen a un mismo resultado compartido antes del impacto (Nesta, 2013; OECD, 2025). En los diagramas, la articulación se representa mediante flechas de cruce entre cadenas paralelas (policy mix), destacando los nodos de convergencia que anticipan los efectos sistémicos.

En la literatura sobre políticas de innovación y transiciones, el *policy mix* se entiende como la combinación de objetivos, instrumentos y procesos de política, y cómo estos se refuerzan o tensionan entre sí (Rogge & Reichardt, 2016; Flanagan, Uyarra & Laranja, 2011). En esta guía, la articulación de teorías de cambio se vincula con esa idea: hacer visibles las interacciones entre intervenciones dentro del sistema de políticas para evaluar no solo resultados de instrumentos aislados, sino la consistencia y coherencia del conjunto en relación con las misiones y transiciones priorizadas en CTCI.

DIAGRAMA 5: ARTICULACIÓN EN LA TEORÍA DE CAMBIO



En la práctica, anidar y articular son dos momentos complementarios. Primero se anida para asegurar que cada instrumento se oriente a los propósitos comunes y sus resultados se sumen coherentemente al nivel superior; luego se articula, conectando esfuerzos intermedios para aprovechar sinergias y evitar duplicidades (OECD, 2025). El anidamiento da estructura y sentido; la articulación otorga dinamismo y colaboración dentro del sistema (BEIS, 2020).

1.3 Temporalidad y plan de evaluación

Evalúa lo observable cuando sea razonable observar. Si bien los plazos dependen del instrumento y su alcance, planteamos algunas referencias:

- Corto plazo (0–2 años): implementación y productos; señales de activación de la cadena (participación, prototipos, convenios, redes incipientes).
- Mediano plazo (3–5 años): resultados intermedios. Por ejemplo, capacidades, adopción tecnológica, vínculos ciencia–empresa y encadenamientos.
- Largo plazo (7–10+ años): impactos agregados, como productividad, sostenibilidad y bienestar.

De la TDC a la agenda de evaluación: para cada ruta, fija hitos a 1, 3 y 5 años que informen decisiones (continuar, escalar, rediseñar o cerrar).

- Inicio: evaluación de proceso y monitoreo de outputs.
- Medio término: evaluación formativa sobre resultados intermedios y supuestos/riesgos; ajustes de ruta.
- Cierre/ex post: evaluación sumativa de resultados e impactos (cuando el rezago lo permita) y análisis de contribución.

1.4 Participación y validación

Participar y validar no es un trámite: es probar la Teoría de Cambio, verificar que los mecanismos sean plausibles, alinear la evaluación con decisiones reales y dar legitimidad al proceso. Se activa en momentos clave del ciclo: al inicio (instalar lógica y usos), con el primer borrador (afinar enfoque y preguntas), a medio camino (corregir desvíos) y cuando cambian premisas, plazos o alcances.

En la práctica, el equipo evaluador analiza la evidencia y propone ajustes; las autoridades priorizan y deciden resguardando independencia y alcance; otros actores, como implementadores, socios, beneficiarios y especialistas, aportan experiencia, datos y señales desde el terreno.

Se recomienda realizar 1 a 2 talleres de Teoría de Cambio para ordenar resultados y supuestos, entrevistas breves donde hay cuellos de botella, y un panel de pares para revisar los puntos metodológicos críticos.

- Para sumar miradas sin reuniones extensas, se recomienda habilitar un canal simple de aportes por escrito y con plazo acotado.
- Registro y memoria del proceso: Para no perder el hilo, se llevan actas de una página por instancia: qué cambió, por qué, evidencia que lo respalda y quién lo aprobó (con fecha). Con un repositorio compartido, esa trazabilidad queda disponible sin burocracia extra.

1.5 Actualización y mínimo viable (TMV)

La MVT es un documento de 1–2 páginas que concentra lo esencial y sirve como pauta de revisión. Cada actualización debe reflejarse en la MVT. Se actualiza cuando el contexto o el aprendizaje lo exigen: en diseño, medio término, cierre o ex post; y también ante cambios relevantes (nueva regulación, shocks, escalamiento o alteración de componentes).

Contenido de la MVT (estructura sugerida):

- Problema y propósito. Qué se busca resolver y para qué.
- Ruta de resultados (del output a resultados).
- Supuestos críticos (5–7), actores clave y temporalidad. Lo que debe ocurrir, quién habilita qué y cuándo.
- Indicadores puente (5–10) y brechas. Qué medimos ahora y cómo cerraremos los vacíos.
- Matriz de medición y análisis (1 pág.) consigna por fila:
 - ➔ Resultado
 - ➔ Pregunta evaluativa
 - ➔ Indicadores y fuentes (dueño, acceso, periodicidad)
 - ➔ Método de estimación
 - ➔ Ventana de decisión (cuándo se usa)
- Bitácora de cambios. Qué cambió, por qué y con qué evidencia.

Se recomienda la actualización mediante un proceso participativo que combine prueba de realidad y teoría negativa (qué podría fallar), dejando registro sistemático de versión, fecha, cambios y evidencia, asignando responsables de aprobación y comunicando la nueva versión a los equipos y autoridades pertinentes.

Articulación entre TDC y el sistema de datos de M&E

Antes de planificar M&E, verifica que la TDC anidada (ver 1.2) esté operativa:

- Resumen TDC y Matriz TDC vigentes.
- Anidamiento según nivel con lectura micro, meso, macro
- Rutas agregables a nivel misión o portafolio e indicadores compartidos definidos.
- Supuestos, riesgos y señales tempranas por ruta, con responsable asignado.
- Hitos de decisión: quién decide y cuándo.

Si falta algo, activa TMV (TDC mínima viable) y MVD (datos mínimos viables) para no detener el proceso, con plan de cierre de brechas.





DATA DE MONITOREO Y EVALUACIÓN

El sistema de Monitoreo y Evaluación (M&E) se diseña desde el uso: qué decisión habilita, quién la toma y cuándo. Debe equilibrar rendición de cuentas y aprendizaje, anclarse a la TDC y aplicar proporcionalidad (esfuerzo según riesgo, escala y novedad).

2.1 Alcance del sistema (anclaje a la TDC)

Antes de medir, define el cambio a observar y la ventana de decisión. Parte de la TDC vigente:

- Nivel de lectura (proyecto, programa o portafolio)
- Horizonte temporal (señales tempranas, resultados, impactos), y
- Desagregaciones relevantes (territorio, sector, tipo de actor, cohorte).

Con esa base, elige la evidencia necesaria según la pregunta. Cada indicador debe vincularse a una decisión concreta.

Fase 1. Planificar (desde la TDC)

Indicadores y periodicidad (qué medir y con qué ritmo). Traduce la TDC en indicadores que gatillen decisiones. Para cada intervención, completa su ficha de propósito y elabora indicadores con ficha íntegra:

- Definición y fórmula
- Universo y desagregaciones
- Fuente, periodicidad y responsable

- Línea base y metas; ventana de decisión, junto con sus límites; y
- Eslabón causal vigilado (A→B) con su supuesto.

Ajusta los ritmos a la maduración: señales tempranas mensuales o trimestrales; impactos en cortes plurianuales o ex post. En misiones o portafolios, agrega pocos indicadores compartidos y explicita cómo cada programa contribuye por rutas.

Fase 2. Captura y gobernanza de datos

Fuentes, acceso e interoperabilidad

Lleva un inventario por indicador: dueño, cobertura, desagregación, costo y condición de uso (pública/restringida/confidencial).

Aplica cinco filtros-umbral de calidad:

- Cobertura y coherencia (integridad del universo, continuidad de series).
- Exactitud (errores, atípicos, sesgos conocidos).
- Oportunidad (a tiempo para la ventana de decisión).
- Comparabilidad (definiciones y desagregaciones consistentes).
- Metadatos completos (metodología, diccionario, versión, contacto).

Si una fuente no supera el umbral, no se integra: documenta el riesgo y activa triangulación como alternativa. Define retención y eliminación (plazos y borrado seguro) y, cuando corresponda, evalúa impacto en privacidad. Minimiza la carga al informante y reutiliza registros administrativos.

Operación con calidad y reproducibilidad

- Procesos y responsabilidades: matriz RACI por fuente/indicador; calendario y puntos de control; comité de datos para cambios metodológicos.
- Flujo de datos y repositorio: captura de datos de forma automatizada cuando sea razonable, procesarlos y dejarlos listos para análisis en repositorio central que reúna datos y scripts, con cuadernos reproducibles (Stata, R, Python), control de versiones y documentación de librerías y entornos utilizados.
- Rigor cualitativo: guías ligadas a preguntas de TDC; muestreo explícito; protocolo de codificación (doble codificación muestral); matriz evidencia-hallazgo; resguardo ético.
- QA/QC³ por etapas:
 - ➔ Antes: estándares, diccionario y prueba básica.
 - ➔ Durante: trazabilidad y verificaciones automáticas.
 - ➔ Después: réplica independiente y nota de robustez.

Fase 3. Analizar e interpretar (de información a evidencia)

- Armoniza unidades y denominadores; unifica definiciones, periodos, códigos e identificadores, y verifica el mínimo de calidad.
- Elabora un plan analítico de una página que conecte pregunta, eslabón de TDC, variables, fuentes, métodos, supuestos y producto.
- Integra con una matriz de triangulación entre cifra, explicación y contexto, y lee el sistema en planos micro, meso y macro.

3 QA (Aseguramiento de la Calidad) y QC (Control de Calidad) son procesos complementarios para garantizar la calidad de productos o servicios.

- Usa técnicas acordes: mapas de actores, análisis de redes, lectura de portafolio, comparaciones y series de tiempo; para mecanismos, Análisis Comparativo Cualitativo (QCA) o rastreo de procesos.
- Interpreta con la regla "si A, entonces B, porque C (bajo C)"; verifica supuestos y factores externos. Si el resultado aún no madura, reporta señales tempranas y su seguimiento.
- Cierra con tres productos verificables: cuaderno reproducible; tablero con series y cortes clave; y fichas de hallazgo con grado de confianza.

Algunos errores frecuentes y cómo evitarlos

- Sobre-atribuir al programa: explicitar factores externos y usar contribución.
- Promediar realidades distintas: segmentar (territorio, tamaño, sector, género) con justificación.
- Ocultar límites de datos: incluir análisis de qué sabemos, qué no sabemos, y qué falta medir.

Fase 4. Comunicar, decidir y mejorar (cerrar el ciclo)

Mantén la trazabilidad desde el dato bruto hasta la conclusión; muestra qué sabemos, qué no sabemos y qué falta medir, y deja los scripts o su versión anonimizada.

La decisión se formaliza con un taller técnico (ajustes operativos), una mesa de decisión (opciones y compromisos) y un acta de adopción que indique qué cambia, quién lo ejecuta, con qué recursos y cómo se medirá el efecto (incluido el ajuste de TDC y del marco lógico).

- Plan B: Si el tiempo es escaso, usa la tarjeta de hallazgo de una página para no perder el momento de decisión, con: pregunta, dato, interpretación, decisión, límites y próximos pasos.

- Mejora continua y MVD. Preguntas guía: ¿lo medido sigue siendo pertinente?, ¿llega a tiempo y en formato útil?, ¿se cumplen protocolos (frecuencia, calidad, trazabilidad, resguardos)?

2.2 Periodicidad y ritmo de actualización

- Anual: cartera de indicadores y definiciones tales como eliminar, mantener, añadir.
- Intermedia: cambios en supuestos de TDC, fuentes o contexto (norma, shock, rediseño).
- Ex post: ajustes tras evaluación final como conservar comparabilidad, documentar cambios.

2.3 Gestión de cambios

Toda modificación se registra de forma trazable: qué se ajustó, desde cuándo y por qué. Aplica control de versiones y nota metodológica para no reescribir lo ya trabajado.

Innovar con criterio: Prioriza automatización en la extracción, transformación y carga para validaciones y tableros para liberar análisis, forma de reducir errores. Antes de sumar fuentes, verifica utilidad para la decisión y cumplimiento de privacidad. Todo cambio se prueba con pilotos acotados y una prueba básica en punto de control para evaluar costo-utilidad y riesgos. Solo con evidencia de beneficio neto avanzamos a adopción plena.

Data Mínima Viable (MVD)

- 1 a 2 páginas con: problema, propósito y resultados.
- 5–7 supuestos críticos, actores clave y temporalidad.
- 5 a 10 indicadores puente y plan para cerrar brechas.
- Matriz compacta de: resultado, pregunta, indicador y su fuente, responsable.
- Bitácora de cambios y plan analítico (de 1 página) que vincule: pregunta, eslabón de TDC, indicador con su fuente, método.





DISEÑO DE LA EVALUACIÓN SISTÉMICA

3.1 Formulación de objetivos y preguntas de evaluación

La formulación de preguntas y objetivos de evaluación se ve influenciada por tres factores principales, los conocimientos o creencias de los actores involucrados, los deseos que tienen o lo que quieren lograr⁴ y las oportunidades u obstáculos que existen (o las limitaciones)⁵ (Befani, 2020). Estos factores se aplican tanto para quienes diseñan y ejecutan las evaluaciones, como para los otros actores interesados, como quienes toman las decisiones o los actores beneficiarios o interesados.

Así, la definición de las preguntas y los objetivos es un proceso articulado con la definición del enfoque y método, que se analizará más adelante.

A modo de orientación, se pueden establecer tres tipos generales de preguntas:

1. Preguntas sobre CUÁNTO o QUÉ TANTO, que se podrían formular de la siguiente manera: ¿Cuánto efecto/impacto produjo la intervención? Que en general es la pregunta por el efecto neto.
2. Preguntas sobre QUÉ efecto se produjo, para QUIÉN y en QUÉ circunstancias o contextos, las cuales no buscan solo un efecto o impacto promedio general, sino dar cuenta de distintos subgrupos de resultados.

⁴ Dicho de otra forma es lo que les gustaría conocer o descubrir a partir de la evaluación, que se relaciona con el uso esperado que tienen para los resultados.

⁵ Estas son todas las condiciones, considerando el tipo de intervención o programa, el tipo de beneficiarios, el contexto en el que se aplica, así como también los recursos pecuniarios, de capacidades y tiempo que se tiene para hacer la evaluación.

3. Preguntas sobre CÓMO se produjo el efecto/ impacto de la intervención o qué proceso o mecanismo (el CÓMO) permitió que la intervención contribuyera a los resultados observados

Esta última pregunta, tiene un subtipo, que se utiliza para poner el foco en si la intervención se comportaría de la misma forma en otros contextos, circunstancias o magnitudes, la cual se formula generalmente como: ¿Qué otros factores deben estar presentes en conjunto con la intervención para producir los resultados observados?

A cada tipo de pregunta se le asocian enfoques y métodos que las responden mejor que otros. Cada evaluación debe elegir una o más preguntas, según lo que se quiera evaluar o conocer por parte de los actores y las condiciones o recursos de que se disponga.

3.2 Requerimientos de información de hallazgos/resultados para la toma de decisión

Se recomienda realizar una matriz que contenga al menos el siguiente listado de actores posibles:

Tipo de actor	Descripción	Se recomienda especificar:
Autoridades y tomadores de decisión	Autoridades de gobierno y directivos que requieren evidencia para la toma de decisión informada para definir prioridades y asignar recursos. Es recomendable especificar los momentos del proceso de la evaluación, en donde serán partícipes	Rol y grado de participación en la evaluación. En otras palabras: qué se espera de ellos. También es recomendable especificar los momentos del proceso de la evaluación, en donde serán partícipes.
Equipos técnicos y de implementación	Ministerios, agencias e instituciones ejecutoras que buscan retroalimentación para mejorar la gestión, coordinar programas y optimizar instrumentos	Posible uso de los resultados y las recomendaciones.
Actores del sistema CTCl	<ul style="list-style-type: none"> • Instituciones de Educación Superior • Centros de Investigación desarrollo e innovación • Empresas • Instituciones Privadas sin Fines de Lucro 	
Conectores	Oficinas de Transferencia y Licenciamiento (OTL), Hubs de transferencia tecnológica, Incubadoras, Aceleradoras, Capital de riesgo y los Nodos CTCl	
Entidades de control y fiscalización gasto público	Instancias como la Dipres y la Agencia de Calidad de las Políticas Públicas.	
Sociedad civil y público general	Organizaciones de la sociedad civil, pueblos originarios, los medios de comunicación, clubes de ciencia, museos, escuelas, entre otros. Ciudadanía interesada en la transparencia, rendición de cuentas y el impacto de las políticas en el desarrollo sostenible.	

Para cada uno, se recomienda completar el rol y grado de participación que tendrá en el proceso de evaluación y el posible uso de los resultados y las recomendaciones que genere la evaluación.

Para una mejor participación o involucramiento de los actores en la evaluación, deberían establecerse momentos y formas de contacto en distintos momentos del proceso, de manera que cuando se les haga algún requerimiento o se espere de ellos una respuesta, esto tenga la suficiente preparación o antecedentes. De modo general, es recomendable informar a los actores de la realización de la evaluación, sus propósitos y alcances, pudiendo recibir también inquietudes y retroalimentación en este momento. Es recomendable comunicar qué se esperará de ellos y en qué fechas, con una adecuada anticipación. Finalmente, al tener los resultados y conclusiones, generar una nueva instancia para informar de las mismas, buscar cierta validación y retroalimentación, e incentivar el uso de los resultados, de la forma que se espere para cada actor⁶.

En este esquema, los actores se encuentran como grupo genérico, pero para el ejercicio real, todos los actores relevantes, que podrían tener roles y usos distintos de su "grupo" deberían estar individualizados.

Luego de completar la tabla, se debería volver a analizar si la pregunta, método y diseño de la evaluación responde adecuadamente a los roles y uso de los resultados de los actores.

⁶ Es recomendable generar distintos formatos de difusión de los resultados, de manera de hacer más sencilla su comprensión según las características e intereses de cada actor.

3.3 Enfoque y método

Cómo ya se mencionó⁷, la elección de enfoque y método de evaluación es una decisión estrechamente relacionada con las preguntas de evaluación a responder y sobre quiénes y cómo se usarán los resultados. Se relacionarán algunos métodos con cada tipo de pregunta definida anteriormente.

→ **Para las preguntas de CUÁNTO o QUÉ tanto**, los métodos más adecuados son:

Experimentales o cuasi experimentales

- Experimento Aleatorio Controlado
- Variables Instrumentales
- Diferencia en diferencias
- Regresión Discontinua

→ **Para las preguntas de QUÉ efecto se produjo, para QUIÉN y en QUÉ circunstancias**, los métodos más adecuados son:

Experimentales o Cuasiexperimentales

- Experimento Aleatorio Controlado
- Diferencia en diferencias
- Emparejamiento de Puntajes de Propensión
- Método de Control Sintético

Basados en Teoría

- Evaluación realista
- Análisis comparativo Cualitativo
- Análisis de contribución

→ Para preguntas sobre CÓMO se produjo el efecto/impacto de la intervención o qué proceso o mecanismo (el CÓMO) permitió que la intervención contribuyera a los resultados observados, los métodos más adecuados son:

⁷ Para revisar el detalle de cada método, con sus fuentes para profundizar y ejemplos, revisar el apartado "Metodologías de evaluación".

Basados en Teoría

- Evaluación realista
- Rastreo de Procesos
- Actualización Bayesiana
- Análisis de Contribución

→ Para preguntas como: **¿Qué otros factores deben estar presentes en conjunto con la intervención para producir los resultados observados?**, los métodos más adecuados son:

Basados en Teoría

- Análisis comparativo cualitativo
- Evaluación realista
- Rastreo de Procesos
- Actualización Bayesiana

Resumen de requerimientos de datos para selección de métodos experimentales o cuasi experimentales:

- Si los participantes o beneficiarios de la intervención se definieron de forma aleatoria, puede aplicarse el **Experimento Aleatorio Controlado**.
- Si se tienen datos tanto antes como después de la intervención, tanto del grupo intervenido como el grupo de control, puede aplicarse **Diferencias en Diferencias**.
- Si los beneficiarios de la intervención se asignan basados en un puntaje de corte, puede aplicarse la **Regresión Discontinua**.
- Si se tienen datos que permitan predecir la participación en la intervención de los sujetos, tanto para los intervenidos como para el grupo de control, puede aplicarse el **Emparejamiento de puntaje de propensión**
- Si se tienen datos históricos suficientes de grupos comparables, puede aplicarse el **Método de Control Sintético**.
- Si existe un factor externo que se relacione con la posibilidad a ser beneficiado de la intervención, pero que no afecte los resultados o impactos, puede usarse el **Método de Variables Instrumentales**.

3.4 Revisión de pares y expertos independientes

La revisión de pares y expertos en las distintas etapas de la evaluación, tanto de su diseño, como de su implementación y resultados, es esencial para garantizar la independencia, transparencia y mejorar la calidad de la evaluación. Los pares y expertos revisores deben ser idóneos tanto para el tema o área de la política a evaluar, como de la metodología que se decide utilizar.

Es necesario definir todos los momentos en donde el proceso de evaluación se someterá a revisión y validación de pares y expertos. Se sugiere considerar al menos para el diseño general de la evaluación, el método elegido y el análisis de los resultados y recomendaciones.

A modo de ejemplo, en el caso de Reino Unido, se cuenta con un grupo de pares revisores externo (Peer Review Group), el cual persigue el objetivo de fortalecer los aprendizajes sobre las evaluaciones (BEIS, 2020). La generación de una red de revisores pares es un desafío pendiente en nuestro sistema CTCl.



DESARROLLO DE LA EVALUACIÓN SISTÉMICA

Antes de iniciar una evaluación, conviene verificar si es pertinente y factible en el momento. Las siguientes preguntas orientan la decisión y permiten preparar la reunión o el documento que la justifique. No son condiciones excluyentes, sino insumos para definir el alcance y la secuencia.

Preguntas orientadoras: ¿es pertinente evaluar ahora?

1. ¿Existe una ventana de decisión clara en los próximos meses (presupuesto, rediseño, continuidad, escalamiento)?
2. ¿Tenemos una Teoría de Cambio vigente -o al menos una versión mínima viable (TMV)- que sirva como columna vertebral?
3. Disponemos de datos suficientes (o factibles en plazo y costo) para responder las preguntas clave y sostener el seguimiento

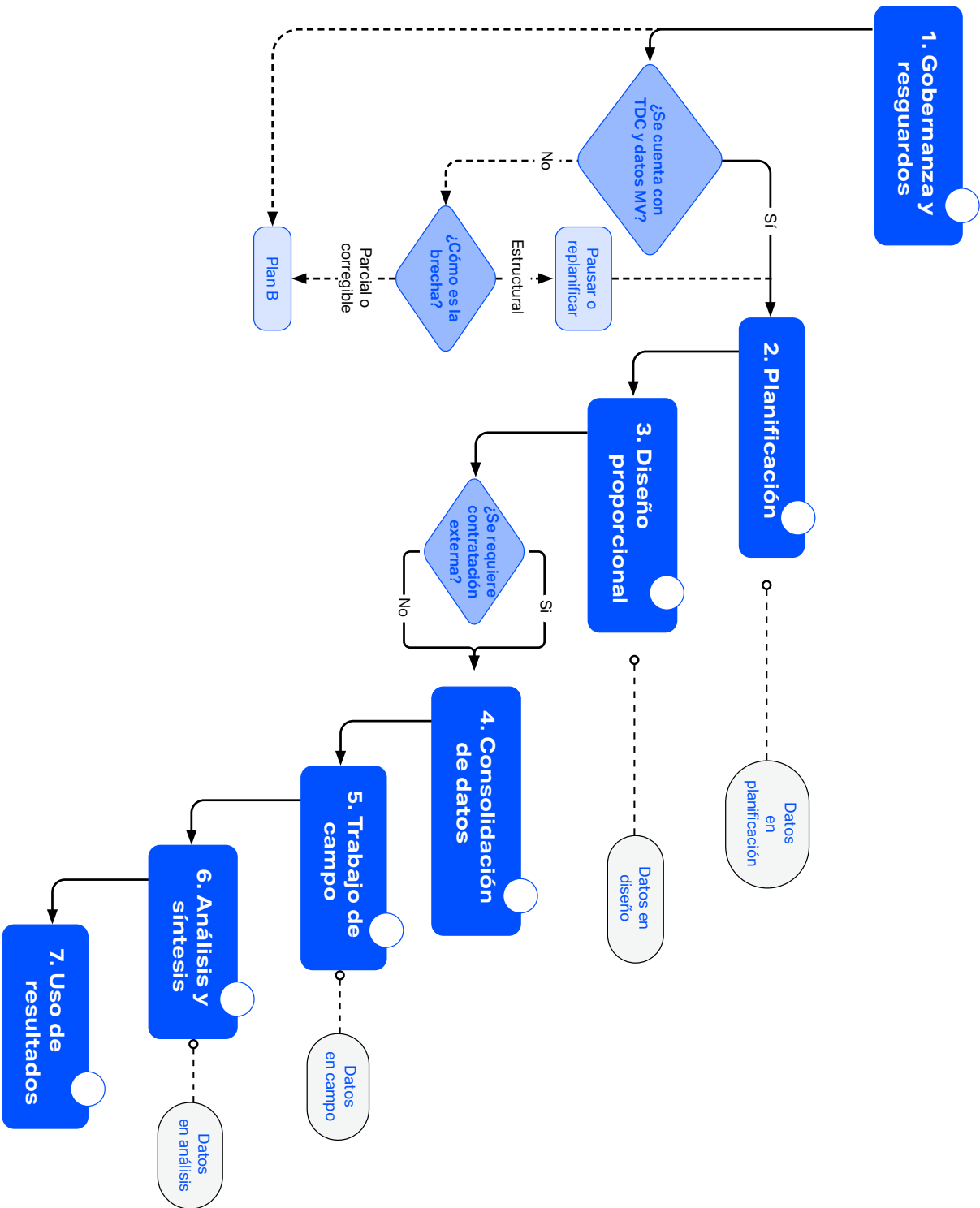
Esta última pregunta se vincula con los principios de gestión de datos y busca anticipar necesidades de información, no detener el proceso.

Si las respuestas anteriores indican que es pertinente continuar (aun cuando algunas se respondan parcialmente), el siguiente diagrama muestra una hoja de ruta operativa para conducir la evaluación desde la preparación hasta el uso de resultados.

El proceso de evaluación sistémica se concibe como un ciclo continuo que avanza por etapas sucesivas, manteniendo activa la trazabilidad de los datos. El Diagrama 6 resume este flujo operativo: desde la verificación de condiciones mínimas (TDC y datos MV) hasta el uso de resultados para retroalimentar la planificación siguiente.

El flujo distingue siete etapas interconectadas. Si no se dispone de TDC o datos mínimos viables, el proceso se pausa o replantea mediante el Plan B; si las brechas son parciales, se activan medidas correctivas proporcionales. El principio de proporcionalidad orienta el esfuerzo en función del riesgo y la escala, mientras que la etapa de uso de resultados retroalimenta la planificación siguiente, consolidando el carácter cíclico del enfoque sistémico.

DIAGRAMA 6: FLUJO OPERATIVO PARA UNA EVALUACIÓN CON ENFOQUE SISTÉMICO



Fuente: Elaboración propia

4.1 Gobernanza y resguardos

Modalidad y capacidades del equipo

La modalidad (interna, externa o mixta) se define según la naturaleza de la brecha detectada y el nivel de independencia técnica requerido. Cuando la brecha sea principalmente de información o método, puede optarse por un equipo interno reforzado; cuando involucra potenciales conflictos de interés o exige juicio independiente, conviene un encargo externo o mixto. En todos los casos, el principio es equilibrar competencias técnicas con resguardo de credibilidad y trazabilidad de resultados.

La decisión de evaluar se integra desde la planificación del programa y se expresa en un mandato formal con atribuciones efectivas. Decidir entre modalidades internas, externas o mixtas, supone la definición del equipo evaluador, considerando lo siguiente.

- Cuando se opte por modalidad interna, conviene equilibrar equipos internos fortalecidos con apoyo externo especializado.
- En modalidades externa o mixta, aplicar reglas espejo de calidad y control para todas las vías.
- Se recomienda establecer protocolos de roles, responsabilidades y coordinación entre miembros del equipo.

Para tomar la decisión, se deben considerar las competencias y disponibilidad de los equipos internos, el tiempo requerido y los recursos disponibles tanto para la ejecución directa como para la contratación del servicio.

Independencia técnica

Instalar mecanismos de resguardo tales como paneles externos (al estilo DIPRES) o comités de revisión con actas públicas y dictámenes técnicos sobre cambios metodológicos relevantes. Se recomienda anticipar canales de escalamiento cuando se identifique bloqueo de datos o permisos relevantes, detallando quién resuelve y en qué plazo.

Resguardos

Los resguardos buscan asegurar que la evaluación cuente desde el inicio con las condiciones mínimas para ejecutarse y con un Plan B diseñado anticipadamente. Este plan define criterios de activación claros, responsables designados y acciones correctivas proporcionales al tipo de brecha.

La evaluación avanza cuando existe un mínimo viable (MV) que garantice trazabilidad y factibilidad técnica:

- Teoría de Cambio breve: con objetivos, supuestos críticos, resultados esperados y dependencias identificadas.
- Datos: inventario de fuentes, accesos y permisos en curso, variables esenciales y metadatos.

Ante esta evaluación, puede ocurrir que sí exista un mínimo viable, o no. En el primer caso, correspondería seguir el proceso según el plan. En el segundo, cabe analizar la naturaleza de la(s) brecha(s) y situación:

- Brecha estructural: pausar o replanificar la evaluación, gestionando las condiciones habilitantes necesarias antes de continuar.

- Brecha parcial o corregible: realizar un ajuste menor(por ejemplo: cambiar fuentes, ajustar muestra o reforzar capacidades), activando el diseño de Plan B. Este debiera incluir riesgos priorizados, acciones correctivas, responsables y tiempos de respuesta. Al activar el Plan B, se ajusta el alcance y/o calendario según la magnitud de la brecha.

Estas orientaciones deben leerse a la luz de los principios en gestión de datos, ya que la trazabilidad, calidad y disponibilidad de la información constituyen resguardos esenciales para la independencia técnica y la credibilidad de la evaluación.

4.2 Planificación

Planificación y coordinación integral

El plan de trabajo cubre todo el ciclo: desde la contratación o conformación del equipo hasta la entrega de productos y su seguimiento. Para alinear la evaluación con ventanas de decisión, la planificación debiera incluir:

- Cronograma, hitos y puntos de control realistas. Considerando en detalle la magnitud del trabajo de campo, en caso de existir.
- Responsabilidades explícitas y coordinación con gestión, monitoreo y política.
- Hitos de decisión: Define hitos que permitan decidir a tiempo si corresponde activar el Plan B.

Gestión administrativa

- Revisar y ajustar tempranamente los procesos administrativos y de licitación para evitar que el proveedor modifique objetivos, enfoques, métodos o plazos fijados en el diseño. Los ajustes permitidos se restringen a aspectos operativos justificados, sin alterar alcance ni estándares de calidad.
- Elabora un inventario inicial de fuentes y se realiza una verificación temprana de acceso a datos (ingreso y descarga) antes de fijar las fechas de trabajo en terreno.

4.3 Diseño proporcional

La magnitud y alcance de la evaluación se deben definir en función del principio de proporcionalidad, considerando al menos importancia del programa a evaluar, magnitud del mismo, riesgo y evidencia anterior. Se consolidan preguntas trazables a la TDC y segmentaciones (territorio, tamaño, tipo de actor, género) para evitar promedios engañosos.

Conformidad con el plan de trabajo.

Para cada pregunta, define: TDC asociada, dato y fuente, cómo se medirá, quién lo hace y qué se entrega. Déjala cerrada antes de salir a terreno; cambios grandes solo con acuerdo del comité.

Decisión operativa: ¿se requiere contratación externa?

La decisión se toma considerando capacidades disponibles, independencia exigida y logística de acceso.

→ Si la respuesta es **SÍ**, considera en las condiciones contractuales o TDR:

- Objetivo, alcance y métodos: Ratificar lo definido en el punto 3. Dejarlo como anexo vinculante.
- Definir entregables y estándares de datos con formato y trazabilidad.
- Garantizar accesos y plazos: responsables de permisos y cronograma con hitos.
- Cambios sustantivos sólo con aprobación escrita del comité.
- Definir criterios objetivos que activen el Plan B cuando se verifique incumplimiento de plazos, falta de acceso a datos o calidad insuficiente de la evidencia.

→ Si la respuesta es **NO**, un acto administrativo interno fijará estándares de calidad equivalentes a los de la contratación externa e instancias de revisión externa.

4.4 Consolidación y estandarización de datos

Aquí se estandarizan metadatos y diccionarios, se fijan identificadores comunes y reglas para cruzar bases; se suscriben acuerdos de intercambio y se establece resguardo ético.

Gobernanza de datos

Además del Plan de Datos Mínimo Viable, que precisa fuentes, permisos, periodicidad, responsables e interoperabilidad, se realiza una prueba de integración que conecta al menos dos fuentes reales con sus metadatos e identificadores.

4.5 Trabajo de campo: Ejecución conforme a diseño metodológico.

El trabajo de campo se organiza según la o las técnicas de investigación a aplicar, tanto cuantitativas (como encuestas) o cualitativas (grupos focales, entrevistas). La recolección se implementa según el punto 3, con control de calidad por fases: captura, depuración, triangulación y validación con actores/pares. Mantener bitácora de decisiones (fecha, motivo, efecto) y un canal técnico con contrapartes para destrabar accesos y validar interpretaciones, sin afectar la independencia del equipo.

Algunos aspectos mínimos de consideraciones éticas son el asegurar consentimiento informado, la minimización de riesgos y la protección de datos personales.

4.6 Análisis y síntesis

Integración evidencia mixta, con lectura en distintos niveles y en las diferentes series de tiempo. Distinguir señales tempranas (outputs), reconfiguraciones (outcomes) e impactos entendidos como contribuciones del portafolio o del ecosistema. Evitar atribuciones lineales cuando la evidencia solo permite inferir contribución.

4.7

Uso de resultados: Recomendaciones y difusión

Implementación y seguimiento de recomendaciones

Acordar con las contrapartes los mecanismos para traducir las recomendaciones en planes de acción concretos. Para ello se recomienda:

- Establecer un sistema de seguimiento estructurado para verificar avances en la implementación.
- Promover la coordinación entre actores cuando las recomendaciones involucren cambios interinstitucionales o multinivel.
- Definir criterios claros sobre quién implementa qué recomendación y cómo se asegura el seguimiento compartido.

Uso y difusión de resultados

Si bien el principal objetivo de las evaluaciones de programas es tomar decisiones de mejora o modificación de tales programas, también el objetivo de transparentar y hacer escrutinio del uso de los recursos públicos es de suma importancia. El carácter público de los resultados de las evaluaciones es entonces esencial.

Además, una evaluación con enfoque sistémico, también busca involucrar más a los actores CTCl y generar una legitimidad de la misma tanto en la ciudadanía como en quienes toman decisiones sobre las políticas del área.

- La comunidad de evaluadores de políticas y académica relacionada a la evaluación también es de relevancia para difundir aprendizajes y generar capacidades de evaluación en el ecosistema CTCl.
- Debe definirse entonces a qué actores se busca llegar y los formatos de presentación de la información más adecuados, como artículos académicos, infografías, webinars, entre otros.



REFLEXIONES SOBRE APLICACIÓN, SOSTENIBILIDAD Y MEJORA CONTINUA

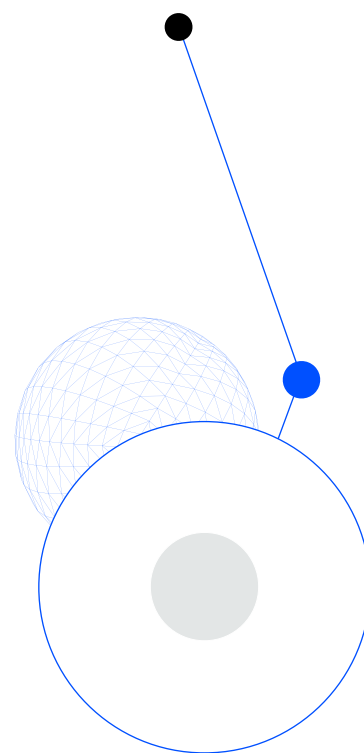
La evaluación sistémica en CTCl requiere combinar orientaciones metodológicas claras con mecanismos de sostenibilidad institucional. La guía se plantea como un insumo operativo que facilita la usabilidad inmediata de las evaluaciones, al vincular teoría de cambio, datos y decisiones de política. No obstante, su proyección dependerá de que el ecosistema incorpore estas prácticas como parte de su funcionamiento habitual.

El uso de esta guía permitirá incorporar un “enfoque sistémico” a las evaluaciones, sin embargo, es necesario fortalecer el avance hacia una concepción sistémica en el sistema de evaluación de la CTCl por medio de medidas posteriores. Durante el proceso de elaboración de este documento, en los talleres intersectoriales así como desde la expertise del equipo consultor surgieron las siguientes sugerencias para continuar con el proceso:

- Mantener la coordinación interinstitucional: Dando proyección a un espacio similar para retroalimentar la elaboración de esta guía. Se sugiere dar formalidad institucional a la constitución de tal espacio.
- Generación de acuerdos sobre estándares: En particular sobre los formatos de los datos, de manera que sean de calidad, seguros y utilizables por todos los actores pertinentes.
- Fortalecer el sistema de datos y monitoreo, avanzando hacia un sistema homologado que permita integrar información de diferentes fuentes y facilitar la evaluación a nivel de programas, instituciones y políticas. Sin un sistema de datos que converse entre actores, la evaluación sistémica difícilmente puede sostenerse en el tiempo.
- Apropiación de la guía: Realización de pilotos de aplicación de la guía, mecanismos de acompañamiento a los equipos técnicos para su utilización en las evaluaciones, y acciones de difusión de la misma.

- Capacitación técnica: Instancias de capacitación en los elementos teóricos asociados a la concepción sistémica de la evaluación, de las metodologías potencialmente aplicables y del sistema de datos y monitoreo. Otra sugerencia en esta línea es conformar una red de evaluación sistémica en CTCl, para intercambiar experiencias, aprendizaje y realizar revisiones y validaciones cruzadas.
- Seguimiento y actualización: Establecimiento de hitos de seguimiento, evaluación y actualización de la misma guía y su nivel de uso e implementación.

En suma, el proceso de elaboración fue ampliamente reconocido por su carácter participativo e interinstitucional, percibiendo la guía como un “puntapié inicial” para transitar hacia evaluaciones con enfoque sistémico.



REFERENCIAS

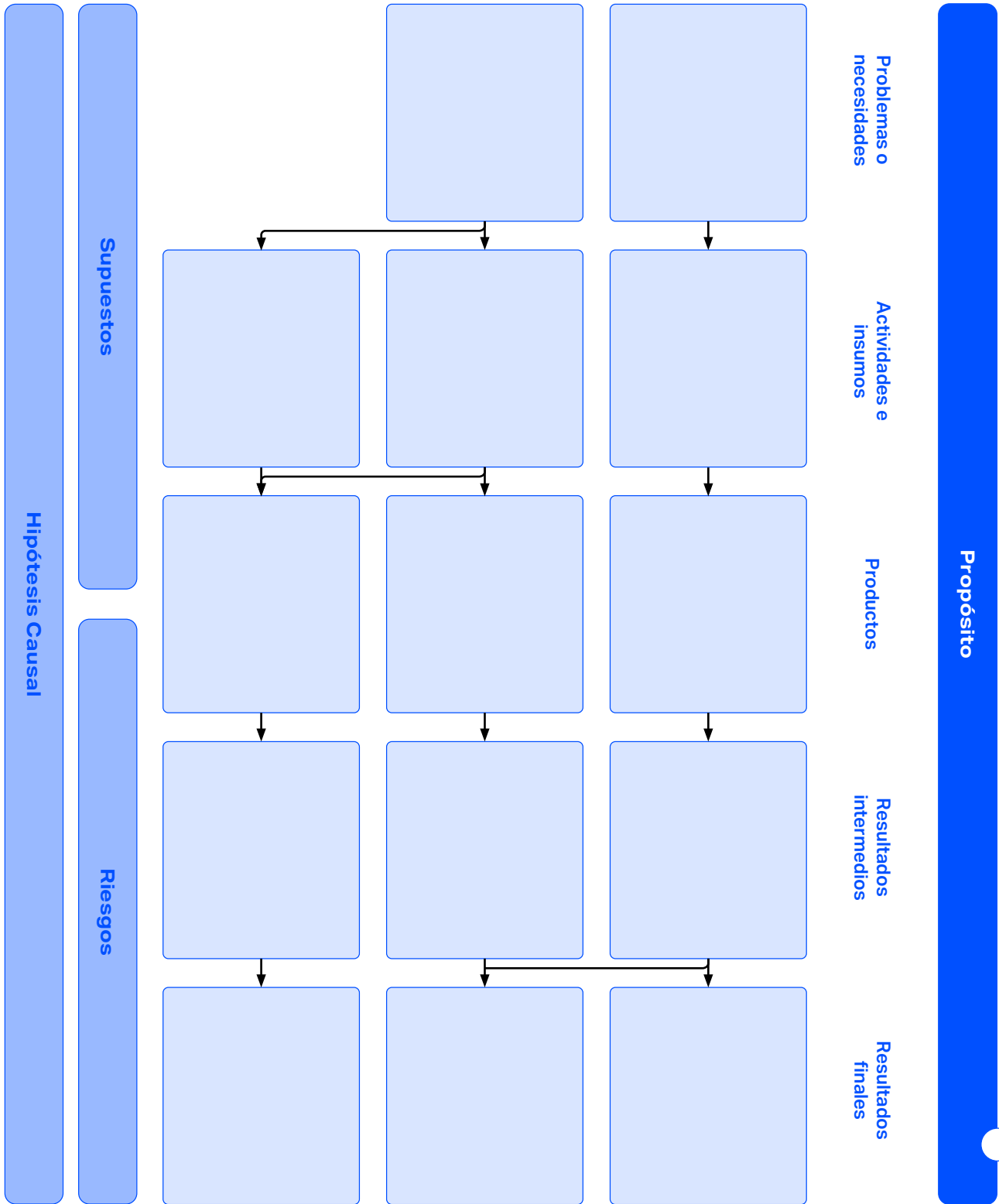
- Arnold, E. (2004). Evaluating research and innovation policy: A systems world needs systems evaluations. *Research Evaluation*, 13(1), 3–17.
- Arnold, E., Gagliardi, D., Giarracca, F., & Simmonds, P. (2018). Impact of Framework Programme research and innovation on European society and economy: A synthesis report. Technopolis Group.
- Braun, D. (2003). Lasting tensions in research policy-making: A delegation problem. *Science and Public Policy*, 30(5), 309–321.
- Bukstein, D., Hernández, E., Monteiro, L., Peralta, M., Reyes, C. y Usher, X. (2020). Evaluación de los programas de innovación empresarial de ANII, 2009-2018. Montevideo: Agencia Nacional de Innovación e Investigación. Disponible en: https://www.anii.org.uy/upcms/files/listado-documentos/documentos/1607621354_informe-de-evaluaci-n-innovaci-n-empresarial.pdf
- Bush, V. (1945). *Science, the Endless Frontier: A Report to the President on a Program for Postwar Scientific Research*. United States Government Printing Office.
- Consejo CTCI, 'Estrategia Nacional de Ciencia, Tecnología, Conocimiento e Innovación para el Desarrollo de Chile - 2022', Consejo de CTCI, Santiago, Chile, Estrategia Nacional de CTCI, 2022.
- Department for Business, Energy & Industrial Strategy. (2020). Monitoring and Evaluation Framework (BEIS Research Paper No. 2020/016). Disponible en: <https://assets.publishing.service.gov.uk/media/5fe341ad-8fa8f56af2a85fcf/beis-monitoring-evaluation-framework.pdf>
- Derrick, G. E., Samuel, G., van Leeuwen, T. N., & Ling, M. (2018). Not all impacts are equal: The grim impact of AI research. *Research Evaluation*, 27(3), 204–214.
- Dipres (2022). Avances del sistema de monitoreo y evaluación y su contribución a la presupuestación. https://www.dipres.gob.cl/598/articles-266615_doc_pdf.pdf
- HM Treasury. (2020). Magenta Book: Central Government guidance on evaluation. HM Government. https://assets.publishing.service.gov.uk/media/5e96cab9d3bf7f412b2264b1/HMT_Magenta_Book.pdf
- Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., ... & Wells, P. (2019). An agenda for sustainability transitions research: State of the art and future directions. *Environmental Innovation and Societal Transitions*, 31, 1–32.
- Larrue, P. (2021). The design and implementation of mission-oriented innovation policies: A new systemic policy approach to address societal challenges. *OECD Science, Technology and Industry Policy Papers*, No. 100, OECD Publishing. <https://doi.org/10.1787/3f6c76a4-en>

- Mazzucato, M. (2018). Mission-oriented research & innovation in the European Union: A problem-solving approach to fuel innovation-led growth. European Commission. https://ec.europa.eu/info/sites/default/files/mazzucato_report_2018.pdf
- Menéndez, M.J., y Villarroel, K. (2024). Análisis del Ecosistema CTCl en Chile en un contexto de Grandes Transformaciones. Consejo Nacional de Ciencia, Tecnología, Conocimiento e Innovación. Santiago, Chile.
- NESTA (2025). Test-and-learn portfolios for systemic change. <https://www.nesta.org.uk/blog/test-and-learn-portfolios-for-systemic-change>
- OCDE, (2020). Chile: Review of Dipres Programme Evaluation System. Disponible: https://www.dipres.gob.cl/598/articles-266584_doc_pdf.pdf
- OCDE, (2025). Proactive portfolio management in mission-oriented innovation policy. Disponible en https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/03/proactive-portfolio-management-in-mission-oriented-innovation-policy_6bc6df0e/1ee956a5-en.pdf
- Patton, M. Q. (2010). Developmental evaluation: Applying complexity concepts to enhance innovation and use. Guilford Press.
- Rosenberg, C., Obando, C. y El-Rayyes, T. (2024). What methods work for evaluating the impact of public investments in RD&I. Department of Science, Innovation and Technology. United Kingdom. Disponible en: <https://www.gov.uk/government/publications/what-methods-work-for-evaluating-the-impact-of-public-investments-in-research-development-and-innovation>
- Technopolis (2024). Consultoría para generar recomendaciones y capacidades para la evaluación sistémica de políticas de ciencia, tecnología, conocimiento e innovación. Disponible en: <https://docs.consejoctci.cl/pdfs.php?i=218>

MATERIAL DE APOYO

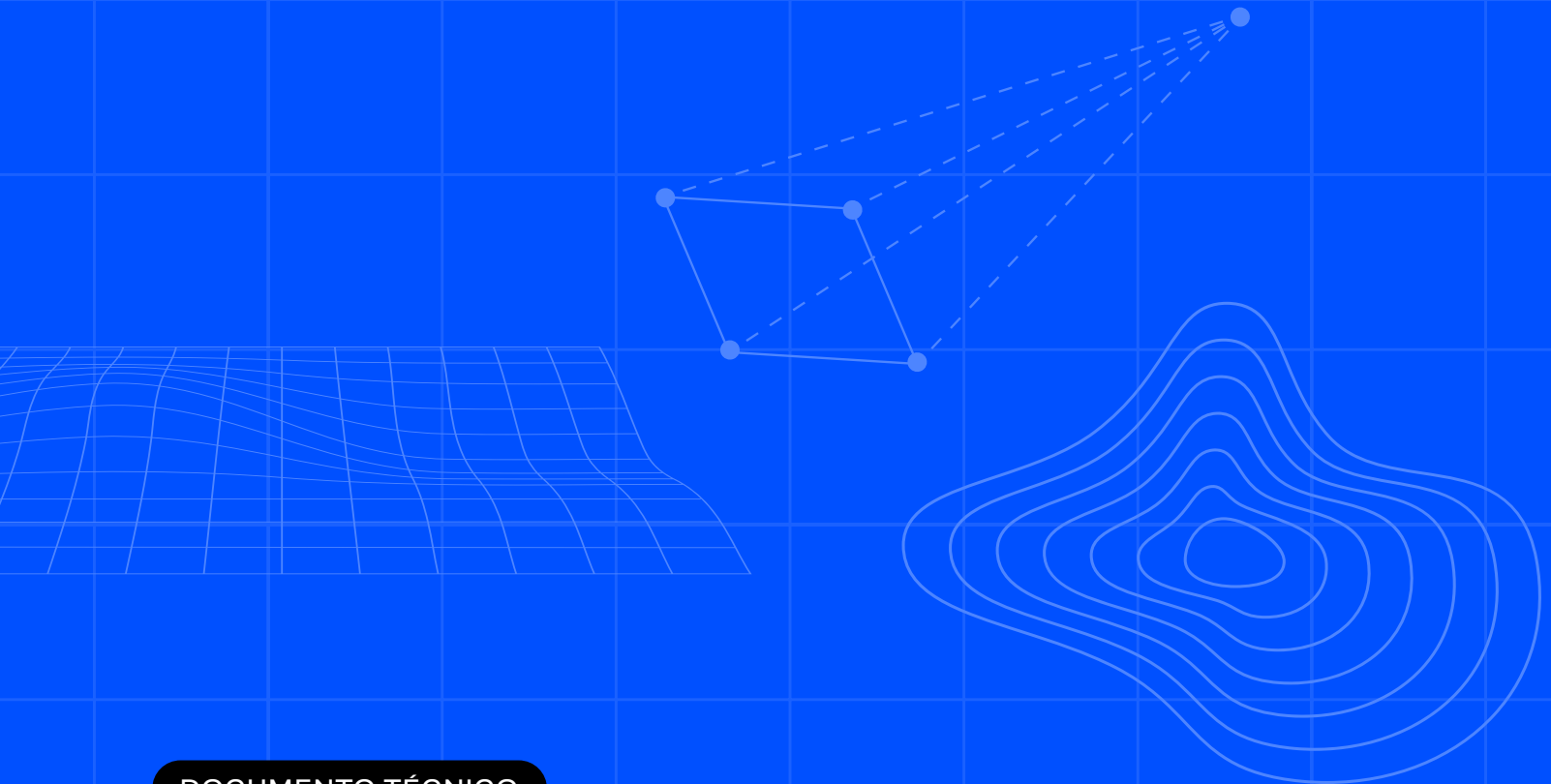
The page features a vibrant blue background with a subtle grid pattern. Several decorative elements are scattered across the page: a series of overlapping circles in the upper right, a large solid circle on the left, and various wavy, organic lines in the bottom left and right corners. A small black horizontal bar is positioned below the main title.

TEORÍA DE CAMBIO





CONSEJO NACIONAL
DE CIENCIA, TECNOLOGÍA,
CONOCIMIENTO E INNOVACIÓN
PARA EL DESARROLLO



DOCUMENTO TÉCNICO

Guía metodológica para la evaluación con enfoque sistémico de Políticas CTCl



CHILE, DICIEMBRE 2025