



Independent
Advisory and
Evaluation
Service

Aplicación del Marco de la QoR4D del CGIAR a las evaluaciones de procesos y desempeño

GUÍAS DE EVALUACIÓN DEL CGIAR 2022

Cita

Independent Advisory and Evaluation Service (IAES). (2022). Guías de evaluación del CGIAR. Aplicación del Marco de la QoR4D del CGIAR a las evaluaciones de procesos y desempeño (versión beta). Roma: Independent Advisory and Evaluation Service.

Diseño y maquetación

Khania Curtis Design and Creative Direction

Título de la guía:

Aplicación del Marco de la QoR4D del CGIAR a las evaluaciones de procesos y desempeño (versión beta): Guías de evaluación del CGIAR 2022

Propósito:

Armonizar el enfoque general para evaluar la calidad de la investigación y la ciencia en el CGIAR y contextos similares; facilitar una comprensión común del criterio de Calidad de la Ciencia (QoS) y de los métodos para aplicarlo.

Audiencias:

La audiencia principal son los evaluadores, gerentes de evaluación y comisarios involucrados en la evaluación de la QoS en el CGIAR. La audiencia secundaria son los responsables de proporcionar aportaciones a las evaluaciones en el Sistema CGIAR. Este documento también puede ser de utilidad para los usuarios en otros contextos de investigación agrícola para el desarrollo (AR4D).

Marco y políticas de referencia:

Esta guía apoya el [Marco de evaluación del CGIAR](#) y la [Política de evaluación de CGIAR \(2022\)](#) y debe leerse junto con otras guías relacionadas con la evaluación. El marco y la política se rigen por el [Marco de calidad de investigación para el desarrollo \(QoR4D\) \(2017, revisado en 2020\)](#).

Contacto:

Para consultas y comentarios sobre el aprendizaje de la puesta en marcha y la aplicación de esta Versión beta, comuníquese con la Función de evaluación del Independent Advisory and Evaluation Service (IAES) de CGIAR al IAES-Evaluations@cgiar.org

Reconocimientos

La Función de evaluación del Independent Advisory and Evaluation Service (IAES) codesarrolló esta guía junto con muchas partes interesadas ([Anexo 1](#)). Desde IAES, Svetlana Negroustoueva, jefa de la función de evaluación, y Stefania Sellitti, consultora en evaluación, dirigieron el proceso bajo la dirección general de Allison Grove Smith, directora del IAES. Estas pautas se basan en las experiencias prácticas de expertos en la materia involucrados en el diseño y evaluación de los Términos de referencia de los Programas de investigación del CGIAR, específicamente Paolo Sarfatti, Jillian Lenne y John Morton. Los participantes contribuyeron aportaciones valiosas en la reunión de validación, durante los debates de las Comunidades de práctica [EvalForward](#) y MELCOP; agradecemos a los respectivos organizadores de estos grupos su facilitación de espacios para el proceso de codiseño. Los autores también desean reconocer la revisión por pares realizada por Rachel Bedoiun, Donna Podems, Marina Apgar y las contribuciones de Sara Vaca, consultora de evaluación y visualización. El trabajo también se benefició de las contribuciones de Holger Meinke y Andrew Ash, miembros del ISDC, así como de Zenda Ofir y Guy Poppy, miembros del [Grupo de referencia de evaluación](#) del IAES. Reconocemos además el trabajo de Science Metrix, Elsevier: la [Nota técnica](#) supuso una aportación importante hacia el proceso de reflexión en torno a esta guía. El equipo más amplio de IAES en Roma garantizó la fluidez en la redacción y el diseño, la asistencia técnica y los procesos administrativos relacionados con el desarrollo de esta guía. El documento fue traducido al castellano por Silvia Montero, y revisado por Sara Vaca y Emma Rotondo.

Petición de aportaciones y retroalimentación

Al IAES le gustaría recibir retroalimentación sobre el aprendizaje de la puesta en marcha y la aplicación de esta versión beta de directrices e invita a los usuarios a contactar con IAES/Evaluación (IAES-Evaluation@cgiar.org) como custodio del documento. La próxima versión continuará basándose en las experiencias directas y estándares en evolución de la industria sobre cómo evaluar la calidad de la ciencia para garantizar de la mejor manera posible la eficacia y rentabilidad de las evaluaciones.

Contenido

1	Introducción	1
1.1	Un contexto cambiante para evaluaciones en el CGIAR	1
1.2	Despliegue de criterios de evaluación para evaluar la calidad de la investigación y la ciencia	2
1.3	Justificación y objetivos de las directrices	3
1.4	Audiencia y usuarios	3
2	¿Qué enmarca la evaluación de la calidad de la ciencia en el CGIAR?	4
2.1	Los parámetros de rendición de cuentas y medición	4
2.2	Marco y política de evaluación	5
2.3	Marco de referencia de calidad de investigación para el desarrollo	6
3	Evaluación de la QoS: dimensiones, preguntas y calendarización	7
3.1	Cuatro dimensiones para evaluar la QoS y mapear otros criterios de evaluación	7
3.2	Preguntas clave de evaluación de la QoS	9
3.3	Calendarización de la evaluación de la QoS en el CGIAR	10
3.4	Garantía de calidad	11
4	Evaluación de la QoS: pasos clave	12
	Paso 1: Alcance y preparación para evaluar la QoS	13
	Paso 2: Identificación de las preguntas y subpreguntas clave de la evaluación	14
	Paso 3: Identificación y mapeo de métodos respecto de las preguntas de evaluación	15
	3a: Identificar los métodos para la recopilación de datos	15
	3b: Mapear los métodos y las preguntas de evaluación	16
	Paso 4: Identificación de documentos clave y posibles brechas de datos y evidencia	17
	Paso 5: Recopilación y análisis de datos	19
	5a: Uso de métodos cuantitativos	19
	5b: Uso de métodos cualitativos, documentación necesaria y rúbricas	21
	5c: Análisis y triangulación de datos	22
	Paso 6: Comunicación de los resultados relacionados con la QoS para mejorar la adopción del aprendizaje y recomendaciones	23
5	Bibliografía	24
	Anexo 1: partes interesadas clave para el desarrollo de directrices	26
	Anexo 2: criterios de evaluación del CGIAR	28
	Anexo 3: ejemplos de preguntas y subpreguntas de evaluación por métodos y fuentes de datos	29
	Anexo 4: parámetros y análisis de datos	31
	Anexo 5: mapeo de los criterios de revisión de propuestas y elementos QoS4D frente a los estándares y criterios del marco de evaluación	33
	Anexo 6: matriz de diseño de evaluación, criterios de la QoS	36
	Anexo 7: uso de rúbricas para la evaluación cualitativa de la QoS	40
	Anexo 8: bibliometría - glosario e indicadores prioritarios	44

Cuadros de Texto

Cuadro de texto 1. Siete criterios de evaluación del CGIAR	5
Cuadro de texto 2. Integración de la QA en el proceso de evaluación	11
Cuadro de texto 3. Ejemplos de métodos cuantitativos y cualitativos recomendados para evaluar la QoS	17
Cuadro de texto 4. Lista de verificación para la recopilación de documentos e información	17

Figuras

Figura 1. Dos paradigmas para las evaluaciones de procesos y desempeño en el CGIAR ...	2
Figura 2. Cinco pilares de la Política de evaluación de CGIAR	3
Figura 3. Cuatro dimensiones interrelacionadas de la evaluación de la QoS	7
Figura 4. Tres preguntas de evaluación para poner en práctica los criterios de evaluación de la QoS	9
Figura 5. Cronograma típico para la duración de un proyecto R4D*	10
Figura 6. Oportunidad para la evaluación de cada criterio	10
Figura 7. Seis pasos clave para la evaluación de la QoS	12
Figura 8. Métodos de evaluación para la recopilación y el análisis de datos	16
Figura 9. Lista de activos de datos	18

Tablas

Tabla 1. Tipos de innovaciones	4
Tabla 2. Consideraciones de los criterios de evaluación de QoS, relevancia y eficacia	6
Tabla 3. Ejemplo de mapeo de siete criterios de evaluación en cuatro dimensiones de la QoS	8
Tabla 4. Criterios de evaluación con ejemplos de preguntas de evaluación	14
Tabla 5. Ejemplos de preguntas y subpreguntas de evaluación por dimensiones de la QoS, fuentes de datos y métodos	16
Tabla 6. Menú de métodos cuantitativos sugeridos para la evaluación de las dimensiones de la QoS, con fortalezas y limitaciones	19
Tabla 7. Temas e indicadores de datos cualitativos por dimensión de la QoS, con criterios de evaluación	20
Tabla 8. Menú de métodos cualitativos sugeridos para la evaluación de la QoS, con fortalezas y limitaciones	21
Tabla 9. Criterios para evaluar la calidad de publicaciones revisadas por pares seleccionadas	22
Tabla 10. Ejemplo de matriz de colecta de datos (DCM) para indicadores bibliométricos de alta prioridad	45

Acrónimos y Abreviaturas

Este documento utiliza los acrónimos y las siglas en su versión original en inglés.

AR4D	Agricultural Research for Development	Investigación agrícola para el desarrollo
ARC	Average of Relative Citations	Media de citas relativas
CAS	CGIAR Advisory Services	Servicios de asesoría del CGIAR
CCAFS	Climate Change, Agriculture and Food Security	Cambio climático, agricultura y seguridad alimentaria
CDI	Citation Distribution Index	Índice de distribución de citas
CIFOR	Center for International Forestry Research	Centro para la Investigación Forestal Internacional
CIP	International Potato Center	Centro internacional de la papa
CIRAD	French agricultural research and international cooperation organization	Centro de Cooperación Internacional en Investigación Agrícola para el Desarrollo
CLARISA	CGIAR Level Agricultural Results Interoperable System Architecture	Arquitectura del sistema interoperable de resultados agrícolas a nivel de CGIAR
CoP	Community of Practice	Comunidad de práctica
CRP	CGIAR Research Program (prior to 2022)	Programa de investigación del CGIAR (antes de 2022)
DAC	Development Assistance Committee	Comité de Ayuda al Desarrollo
DCM	Data Collection Matrix	Matriz de colecta de datos
EA	Evaluability Assessment	Análisis de evaluabilidad
EF	Evaluation Framework	Marco de evaluación
EP	Evaluation Policy	Política de evaluación
EQ	Evaluation Question	Pregunta de evaluación
ERG	Evaluation Reference Group	Grupo de referencia de evaluación
FAO	Food and Agriculture Organization of the UN	Organización de las Naciones Unidas para la Alimentación y la Agricultura
FGD	Focus Group Discussion	Grupo de discusión focal
FIA	Fund for International Agricultural Research	Fondo internacional para la investigación agrícola
GDI	Gender, Diversity, and Inclusion	Género, diversidad e inclusión
GENSA	Gender and Equity Network South Asia	Red de género y equidad, Asia meridional
GFAR	Global Forum on Agricultural Research and Innovation	Foro global sobre la investigación e innovación agrícola
GIZ	German Corporation for International Cooperation	Sociedad alemana para la Cooperación Internacional
GLDC	Grain Legumes and Dryland Cereals	Granos, legumbres y cereales
HCP	Highly Cited Publications	Publicaciones altamente citadas

Acrónimos y Abreviaturas cont'd

IAES	Independent Advisory and Evaluation Service	Servicio independiente de asesoramiento y evaluación
ICARDA	International Center for Agricultural Research in the Dry Areas	Centro internacional para la investigación agrícola en zonas áridas
IDRC	International Development Research Centre	Centro internacional de investigaciones para el desarrollo
IEA	Independent Evaluation Arrangement	Independent Evaluation Arrangement
IPG	International Public Goods	Bienes públicos internacionales
IR	Inception Report	Informe inicial
ISDC	Independent Science for Development Council	Consejo independiente de ciencia para el desarrollo
IUCN	International Union for Conservation of Nature	Unión internacional para la conservación de la naturaleza
KII	Key Informant Interviews	Entrevistas con informantes clave
KM	Knowledge Management	Gestión de conocimiento
LMIC	Low and Middle Income Countries	Países de renta media-baja
M&E	Monitoring and Evaluation	Monitoreo y evaluación
MARLO	Managing Agricultural Research for Learning and Outcomes	Sistema de gestión de la investigación agrícola para el aprendizaje y los resultados
MEL/IA	Monitoring, Evaluation and Learning/Impact Assessment	Monitoreo, evaluación y aprendizaje/ evaluación de impacto
MELCOP	Monitoring, Evaluation and Learning Community of Practice	Comunidad de práctica de seguimiento, evaluación y aprendizaje
MR	Management Response	Respuesta de la gerencia
NARS	National Agricultural Research Systems	Sistemas nacionales de investigación agrícola
NRI	Natural Resources Institute	Instituto de recursos naturales
OECD/DAC	Organisation for Economic Co-operation and Development/ Development Assistance Committee	Comité de ayuda al desarrollo de la Organización para la Cooperación y el Desarrollo Económico
OICR	Outcome Impact Case Reports	Informes de casos de impacto de los resultados
PCU	Project Coordination Unit	Unidad de coordinación de proyectos
PIM	Policies, Institutions and Markets	Políticas, instituciones y mercados
PPU	Performance Portfolio Unit	Unidad de desempeño de la cartera
PRMF	Performance and Results Management Framework	Marco de gestión de desempeño y resultados
PRMS	Performance and Results Management System	Sistema de gestión de desempeño y resultados
QA	Quality Assurance	Garantía de calidad

Acrónimos y Abreviaturas cont'd

QoR4D	Quality of Research for Development	Calidad de investigación para el desarrollo
QoS	Quality of Science	Calidad de la ciencia
R4D	Research for Development	Investigación para el desarrollo
R&D	Research & Development	Investigación y desarrollo (I+D)
REF	Research Excellence Framework	Marco de excelencia en investigación
RQ+	Research Quality Plus Assessment Instrument	Instrumento de evaluación de la calidad de investigación plus
RTB	Roots, tubers and bananas	Raíces, tubérculos y bananas
SDGs	Sustainable Development Goals	Objetivos de desarrollo sostenible
SMEs	Subject-matter experts	Expertos en la materia
SNA	Social Network Analysis	Análisis de redes sociales
SSH	Social Sciences and Humanities	Ciencias sociales y humanidades
ToC	Theory of Change	Teoría de cambio
ToR	Terms of Reference	Términos de referencia
TRA	Technical Reporting Arrangement	Technical Reporting Arrangement
UNPFA	United Nations Population Fund	Fondo de población de las Naciones Unidas
WLE	Water, Land and Ecosystems	Agua, tierra y ecosistemas

ESTAS GUÍAS DE EVALUACIÓN TIENEN COMO INTENCIÓN ARMONIZAR EL ENFOQUE GENERAL PARA LA EVALUACIÓN DE LA CALIDAD DE LA INVESTIGACIÓN Y LA CIENCIA EN EL CGIAR. SON NECESARIAS PARA GARANTIZAR UNA COMPRENSIÓN COMÚN DE ESTE CRITERIO DE EVALUACIÓN Y LOS MÉTODOS DISPONIBLES ENTRE LOS EXPERTOS EN LA MATERIA Y LOS EVALUADORES QUE IMPLEMENTAN LAS EVALUACIONES.

COMO TAL, ESTA GUÍA PROPORCIONA UN MENÚ DE MÉTODOS PARA LOS EVALUADORES Y TRANSPARENCIA PARA LOS USUARIOS DE LAS EVALUACIONES, INCLUIDOS LOS PROFESIONALES DE MONITOREO, PARA INTERPRETAR LOS HALLAZGOS Y PONER EN PRÁCTICA EL CRITERIO DE EVALUACIÓN DE LA CALIDAD DE LA CIENCIA (QOS).

Las evaluaciones rigurosas e independientes de procesos y desempeño juegan un papel fundamental para ayudar a CGIAR a informar el diseño y la implementación de sus investigaciones e innovaciones. Las evaluaciones proporcionan evidencia procesable para las decisiones de gestión y gobernanza, facilitan el aprendizaje y aseguran la rendición de cuentas a los financiadores y otros grupos de partes interesadas. El Marco de evaluación del CGIAR (EF) y la Política de evaluación (EP) (2022) articulan cómo las evaluaciones de procesos y desempeño ayudan a CGIAR a cumplir su misión e implementar su Estrategia de investigación e innovación del CGIAR 2030 (Estrategia CGIAR 2030).¹

La actividad principal de CGIAR es brindar investigación para el desarrollo. Esta guía para evaluadores y expertos en la materia explica los impulsores, parámetros y métodos para evaluar la calidad de la ciencia y la investigación de CGIAR.

Esta guía de evaluación se basa en el marco de referencia (2020) de Calidad de investigación para el desarrollo (QoR4D) y fue informada por el Marco de excelencia en investigación (REF) y el instrumento de evaluación de la calidad de investigación plus (RQ+). El Marco y la Política de evaluación del CGIAR (2022) y el Marco de gestión de desempeño y resultados de CGIAR 2022-2030 estableció las bases para esta guía en el contexto de CGIAR.

1.1 Un contexto cambiante para evaluaciones en el CGIAR

El CGIAR es una alianza mundial de investigación para un futuro con seguridad alimentaria cuya misión es brindar ciencia e innovación para transformar los sistemas alimentarios, terrestres e hídricos en una crisis climática. Como parte del proceso de reforma de One CGIAR,² la Estrategia CGIAR 2030 se esfuerza por lograr un cambio transformador en tres áreas de acción (Transformación

de sistemas, Sistemas agroalimentarios resilientes e Innovación genética) y cinco áreas de impacto centradas en los Objetivos de Desarrollo Sostenible (los ODS), llevados a la práctica a través de más de 30 iniciativas regionales y globales.

La Estrategia CGIAR 2030³ utiliza las siguientes definiciones:

Investigación – Generación y comunicación de datos, información y conocimiento con base empírica.

Ciencia – Investigación rigurosa basada en la teoría.

El Marco de gestión de desempeño y resultados (PRMF) apoya la implementación de la Estrategia CGIAR 2030. Proporciona la base para la rendición de cuentas, el aprendizaje, la comunicación y la movilización de recursos de CGIAR, y sirve como base para el Acuerdo de informes técnicos (2022) (Figura 2).

¹ La Política de evaluación del CGIAR debe revisarse periódicamente. El IAES del CGIAR puede recomendar enmiendas según corresponda al Consejo del Sistema y la Junta para su aprobación. Las revisiones futuras pueden, p. ej., tener en cuenta si podría mejorarse la agenda de transformación del sistema de CGIAR con criterios de evaluación adicionales.

² One CGIAR es un proceso emprendido por el CGIAR iniciado en 2019 para fortalecer sus alianzas, conocimientos, activos y presencia global. Tiene como objetivo lograr una mayor integración e impacto frente a los desafíos interdependientes que enfrenta el mundo actual. <https://www.cgiar.org/food-security-impact/one-cgiar/>

³ CGIAR [Estrategia de investigación e innovación de CGIAR 2030](#)

Despliegue de criterios de evaluación para evaluar la calidad de la investigación y la ciencia

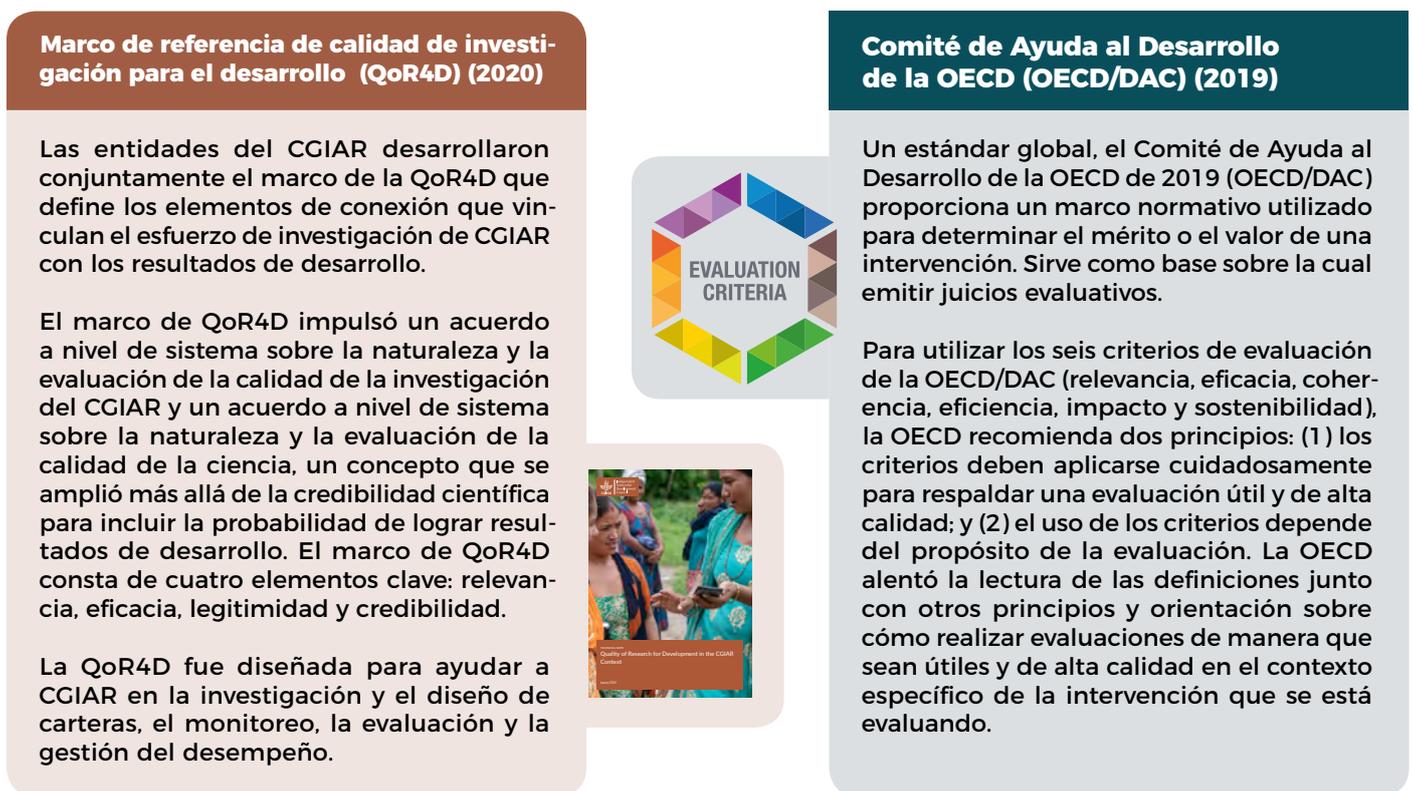
Evaluar la calidad de la investigación y la ciencia del CGIAR es importante por tres razones principales: (1) para rendir cuentas de la inversión pública y privada en investigación que genera bienes públicos internacionales; (2) para informar a los financiadores sobre la calidad tanto de los procesos científicos como de los resultados científicos de la investigación realizada; (3) y para proporcionar evidencia sobre cómo la ciencia del CGIAR contribuye a los objetivos de la organización, incluso como parte de un esfuerzo más amplio, para contribuir al progreso acelerado hacia los ODS.

Para el caso específico de CGIAR, su actividad principal yace en la investigación para el desarrollo (R4D). La Estrategia CGIAR 2030 define esto como investigación y ciencia. Los financiadores de CGIAR se enfocan en los resultados de desarrollo. En consecuencia, la evaluación

de las intervenciones del CGIAR debe responder a los dos marcos de QoR4D (investigación orientada a generar resultados de desarrollo) y OECD/DAC (orientada al desarrollo). Los criterios estándar de evaluación de la ayuda al desarrollo de la OECD/DAC son insuficientes para evaluar la actividad principal del CGIAR. Es necesario tener en cuenta las características especiales de la R4D en la planificación y el enfoque de la evaluación. De forma específica, este documento contempla la naturaleza impredecible y arriesgada de la investigación y la dilatación del tiempo inherente a la materialización de los resultados.

Los métodos y criterios de evaluación del CGIAR responden a dos marcos fundamentales importantes para el CGIAR y sus partes interesadas, que conforman la base para evaluar la calidad de la investigación y la ciencia (Figura 1).

Figura 1: Dos paradigmas para las evaluaciones de procesos y desempeño en el CGIAR



Los siete criterios de evaluación de CGIAR, que responden a los marcos de la QoR4D y de la OECD/DAC, enmarcan y brindan estructura al enfoque sustancial de las preguntas de evaluación (Anexo 2; CGIAR Política de evaluación 2022).

Se ha utilizado un criterio de evaluación de calidad de la ciencia (QoS) en las evaluaciones de CGIAR durante

al menos la última década (consulte la Política de evaluación de 2012).⁴ Guiado por el marco de la QoR4D, se aplicó el criterio de evaluación de la QoS en las Revisiones evaluativas de 2020 de los 12 programas de investigación del CGIAR (los CRP) junto con el criterio de eficacia. Se aplicaron métodos mixtos (cuantitativos y cualitativos) para abordar ambos criterios de evaluación.⁵

⁴ (retirado) Política del CGIAR para la evaluación externa independiente (2012) reemplazada por la versión de 2022.

⁵ Para detalles adicionales, consulte el [Informe y anexos íntegros de los CRP](#) para las revisiones evaluativas individuales de los CRP y/o para leer los cuatro estudios de caso en la nota técnica desarrollada conjuntamente por Science-Metrix y CAS/Evaluación (2022).

1.3 Justificación y objetivos de las directrices

Los dos marcos fundamentales, QoR4D y OECD/DAC, comparten sinergias y complementariedades, así como terminologías y conceptos superpuestos. Estas sinergias y superposiciones proporcionan **la justificación** de esta guía, que pretende aclarar cómo evaluar la investigación y la ciencia que se lleva a cabo para el desarrollo y contribuye hacia los ODS.

Las directrices se enfocan en la puesta en práctica del criterio de evaluación de la QoS y en el despliegue de los siete criterios de evaluación del CGIAR en combinación para la emisión de juicios evaluativos de las actividades de investigación para el desarrollo. La evaluación específica de la QoS es crucial para verificar si la ciencia producida es creíble (robusta) y generada de forma legítima (justa y ética) a través de la cooperación de múltiples partes interesadas y si es coherente con las necesidades de las personas sobre el terreno.

Esta guía para evaluadores, Aplicación del Marco de la QoR4D del CGIAR a las evaluaciones de procesos y desempeño, tiene como objetivos principales:

- facilitar una comprensión común del criterio de evaluación de la QoS, incluso con relación a otros criterios de evaluación;
- resumir un enfoque común para evaluar la investigación y la ciencia y proporcionar un menú de métodos basado en la revisión crítica de sus fortalezas y desafíos;
- aportar una referencia cruzada entre las medidas ex ante de ISDC y las medidas de la Política de evaluación de CGIAR para la evaluación intermedia/ex post;
- destacar las funciones y responsabilidades para facilitar la evaluación de la QoS en CGIAR en diferentes niveles de acuerdo con los principios del Marco de evaluación de «medibilidad» y «rendición de cuentas mutua» y subrayar el Pilar 2 (Enfoque holístico y coherente para el monitoreo y el aprendizaje a nivel de todo el CGIAR) de la Política de evaluación de CGIAR (Figura 2).

Figura 2: Cinco pilares de la Política de evaluación de CGIAR



1.4 Audiencia y usuarios

Estas pautas están destinadas principalmente a evaluadores, gerentes de evaluación y comisarios involucrados en evaluaciones de procesos y desempeño de investigación y ciencia en CGIAR. También serán útiles para los responsables de proporcionar aportaciones a las evaluaciones en el CGIAR System Organization, p. ej., gerentes de investigación y profesionales de Monitoreo, Evaluación y Aprendizaje (MEA).

Los enfoques descritos son lo suficientemente amplios para ser utilizados en muchos contextos, tanto dentro como fuera del CGIAR. Pueden ser adaptados, en caso necesario, por los socios de CGIAR y agencias afines que realizan investigación para el desarrollo.

Familiaridad con el Marco de evaluación del CGIAR y la Política de evaluación es un requisito previo para comprender y utilizar estas directrices.⁶ Las directrices se lanzan al comienzo del ciclo económico de CGIAR y el Acuerdo de informes técnicos (TRA, 2022) para la realización de las evaluaciones de procesos y desempeño de la nueva cartera de CGIAR. Para poder abarcar todos los elementos de la QoR4D, una evaluación de procesos y desempeño de una intervención de R4D debe contar con suficientes recursos para abordar, como mínimo, los criterios de relevancia, eficacia y calidad de la ciencia (legitimidad y credibilidad).⁷

⁶ La Guía «Elaboración y uso de los análisis de evaluabilidad en el CGIAR» (2022) y las Directrices de compromiso y respuesta de la gerencia (en desarrollo) tienen como objetivo dirigir la atención a elementos clave como los recursos y las infraestructuras relacionados con MEA, es decir, los sistemas de gestión de resultados de desempeño.

⁷ Una fuente adicional de aprendizaje sobre la Política y el Marco de Evaluación de CGIAR <https://youtu.be/DeUn7T3UW6s> (video).

¿Qué enmarca la evaluación de la calidad de la ciencia en el CGIAR?

2.1 Los parámetros de rendición de cuentas y medición

La Estrategia de investigación e innovación 2030 enfatiza el papel crítico de la ciencia y la innovación en el suministro de nueva evidencia, conocimientos y soluciones que alimentan las alianzas estratégicas para el cambio. El Marco de gestión de desempeño y resultados 2022-2030 (PRMF) que le sirve como base establece la intención de CGIAR de medir sus contribuciones desde la investigación hasta el impacto a lo largo tres rutas principales: (1) capacitación específica, (2) asesoramiento sobre políticas e (3) innovaciones y tecnologías basadas en la ciencia (Estrategia CGIAR 2030 [p. 20]; Tabla 1). El PRMF, puesto en práctica en parte a través del TRA (2022), guiará e informará la toma de decisiones durante cada etapa en las iniciativas de

investigación durante el ciclo de proyecto; es decir, procesos de seguimiento y aprendizaje para ayudar a los equipos de las iniciativas a gestionar sus programas de forma dinámica durante la ejecución, a través de la validación y/o adaptación basadas en evidencia de las teorías del cambio; para ayudar a la supervisión por parte de gerentes de la implementación de la cartera; y para apoyar las decisiones de inversión.

Por tanto, un papel importante para la evaluación es comprender si los tipos de innovación y otros resultados mencionados anteriormente muestran ciencia de alta calidad y contribuyen a las rutas hacia los resultados de desarrollo, y cómo lo hacen.

Tabla 1: Tipos de innovaciones

Innovaciones: Paquetes de aportaciones complementarias necesarios para desarrollar y escalar productos, servicios y soluciones.

Capacitación 	Política 	Tecnologías 
El saber hacer y las capacidades de individuos, empresas, organizaciones, y redes para diseñar, probar, validar y utilizar innovaciones.	Políticas públicas, legislación, ejecución pública y privada y estrategias comerciales que crean un entorno propicio en el que las innovaciones pueden escalarse, o que en sí mismas representan innovaciones generadoras de impacto.	Las variedades, máquinas, prácticas de gestión, productos y herramientas - incluyendo big data y herramientas de información - cuyo uso puede generar beneficios, ganancias o eficiencias y cuyo despliegue a escala puede generar impacto.

Fuente: PRMF 2022-2030, p. 5

2.2 Marco y política de evaluación

En respuesta al ritmo creciente de cambio para apoyar las necesidades y demandas cambiantes del CGIAR, la Junta del Sistema (23ª sesión) y el Consejo del Sistema (15ª reunión) aprobaron el nuevo Marco de evaluación adecuado para su propósito y la Política de evaluación del CGIAR revisada en 2022. El Marco y la Política definen y establecen 15 estándares y principios, así como el enfoque general para las evaluaciones de procesos y desempeño en el CGIAR.

Evaluaciones de procesos son evaluaciones del funcionamiento organizacional, instrumentos, mecanismos y prácticas de gestión de asuntos institucionales y de procedimiento en todo el CGIAR y evaluaciones de la experiencia con marcos, políticas, criterios y procedimientos del CGIAR.

Evaluaciones de desempeño proporcionan evidencia rigurosa e imparcial de la eficacia organizativa y los modelos operativos mediante la evaluación del progreso hacia el logro de resultados mediante la comparación de datos de desempeño con el objetivo declarado y la presentación de informes según un calendario predeterminado, para informar la toma de decisiones sobre cómo mejor utilizar o invertir recursos financieros o técnicos, resolver desafíos y apoyar el progreso continuo.

Alineados con el marco de la OECD/DAC, los siete criterios de evaluación brindan estructura al enfoque sustancial de las preguntas de evaluación (Cuadro de texto 1 [Anexo 3, Política de evaluación de CGIAR 2022]).

Cuadro de texto 1: *Siete criterios de evaluación del CGIAR (consulte el Anexo 2 para más información)*

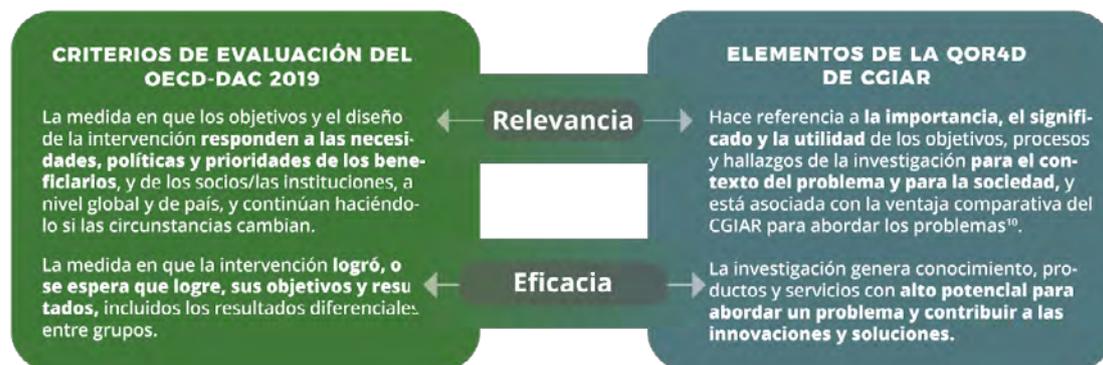
- Relevancia
- Coherencia
- Eficacia
- Eficiencia
- QoS
- Sostenibilidad
- Impacto

El Marco de referencia de calidad de la investigación para el desarrollo (Qo4RD) tiene como objetivo brindar coherencia y mejorar la calidad general de la R4D dentro de la cartera del CGIAR, registrando un acuerdo a nivel de todo el sistema sobre la naturaleza y la evaluación de la calidad de la investigación. Abarca la relevancia para los grupos de usuarios, credibilidad científica, legitimidad y probabilidad de lograr resultados de desarrollo, desglosados en cuatro elementos clave: relevancia, credibilidad científica, legitimidad y eficacia.

En el marco de la QoR4D, relevancia y eficacia se definen de forma ligeramente diferente de los criterios de evaluación de la OECD/DAC. En CGIAR estos dos criterios utilizan los mismos términos y aluden de forma directa a

los elementos de la QoR4D y los criterios de la OECD/DAC (Figura 1). Los criterios de evaluación del CGIAR responden a las definiciones según la QoR4D y están alineados con la OECD/DAC para evaluaciones ex post; p. ej., una evaluación de procesos y desempeño⁸ de eficacia debe tener en cuenta tanto la fidelidad al plan (según la OECD/DAC) como la preparación para el uso (según la QoR4D). El criterio de relevancia para la OECD/DAC se refiere a la capacidad de respuesta al desafío del desarrollo, mientras que en la QoR4D el elemento de relevancia también incorpora una asociación con la ventaja comparativa del CGIAR. La Tabla 2 presenta diferencias en las definiciones de los criterios de relevancia y eficacia entre los dos marcos (consulte también la Figura 1).

Tabla 2: Consideraciones en torno a «relevancia» y «eficacia»⁹



El criterio de evaluación de la QoS asegura la cobertura evaluativa de los otros dos elementos de la QoR4D (legitimidad y credibilidad científica), que no son criterios explícitos de la OECD/DAC, pero son vitales en el marco de la QoR4D (Figura 1). Como uno de los siete criterios, el criterio de evaluación de la QoS refleja la identidad del CGIAR como una alianza mundial de investigación para el desarrollo. Los criterios unificados de evaluación de la QoS se basan en la utilización de métodos mixtos para juzgar el grado y el alcance de la legitimidad y la credibilidad científica.

Un mapeo de los elementos de la QoR4D y los criterios de revisión del ISDC a los estándares y principios de evaluación en el Anexo 5 ilustra la alineación entre los marcos de Evaluación y de la QoR4D, con el propósito específico de realizar evaluaciones de procesos y desempeño en el CGIAR.

⁸ Las definiciones aparecen en la sección 2.2.

⁹ De acuerdo con la Política de evaluación, otros criterios de evaluación (eficiencia, coherencia, sostenibilidad, impacto) son coherentes con las definiciones de la OECD-DAC (consulte el Anexo 2).

¹⁰Independent Science for Development Council. 2022. Identifying and Using CGIAR's Comparative Advantage [Identificación y utilización de la ventaja comparativa del CGIAR]. Roma: CGIAR Independent Advisory and Evaluation Service. <https://iaes.cgiar.org/sites/default/files/pdf/ISDC-Technical-Note-Identifying-and-Using-CGIAR-Comparative-Advantage.pdf>

Evaluación de la QoS: dimensiones, preguntas y calendarización

El contexto y la calendarización de la evaluación y las características del objetivo de la evaluación (la entidad evaluada) dictarán el enfoque general de la evaluación y la selección de métodos. De acuerdo con los tipos de evaluación clave en la Política de Evaluación del CGIAR,

es más probable que las evaluaciones enfocadas de QoS se lleven a cabo como parte de un proceso más amplio y/o una evaluación del desempeño (Secciones 2.2 y 2.3), p. ej., las Revisiones de los CRP 2020 que utilizaron criterios evaluativos de eficacia y de la QoS.¹¹

3.1 Cuatro dimensiones para evaluar la QoS y mapear otros criterios de evaluación

En miras a una evaluación integral de la QoS, esta guía recomienda el análisis de cuatro dimensiones interrelacionadas: diseño de investigación, insumos, procesos y resultados (Figura 3).¹²

- **Diseño de la investigación:** La idoneidad del diseño de investigación implementado se juzga en términos de estándares comúnmente aceptados en un campo temático designado. Evaluación de la adecuación técnica de la agenda y estrategia de investigación, y relevancia y coherencia generales.

- **Insumos:** Los insumos necesarios de la investigación se evalúan en términos de su adecuación con relación a los resultados. Ejemplos de insumos incluyen personal de investigación, composición de los equipos, disponibilidad de infraestructura de investigación y financiación adecuadas.

- **Procesos:** La gestión y la coordinación están impulsadas por incentivos para lograr y mantener la alta credibilidad científica de los resultados. Además, la evaluación explora la eficacia de la construcción y el

aprovechamiento de alianzas, es decir, si son basadas en un entendimiento mutuo, la confianza y el compromiso, con un claro reconocimiento de diversas perspectivas, necesidades, roles y contribuciones. También se evalúa la equidad y los aspectos éticos de la implementación real de la cartera de investigación. Por ejemplo, la medida en que los gerentes de la investigación y los procesos implementaron orientaciones de ética y procesos de toma de decisiones y la representatividad de los mismos.

- **Resultados:** Los resultados variarán de forma considerable; sin embargo, la mayoría de las veces son productos o servicios tangibles. Por ejemplo, nuevas semillas o germoplasma, o productos técnicos como documentos de políticas, artículos de revistas, resúmenes técnicos y nuevas técnicas de gestión del suelo. Otros resultados técnicos incluyen software, directrices, herramientas de apoyo a la toma de decisiones y materiales de capacitación, resúmenes de políticas y otras acciones orientadas al cambio de políticas.



Figura 3: Cuatro dimensiones interrelacionadas de la evaluación de la QoS

¹¹Términos de referencia (y apéndice) para las Revisiones de los CRP 2020

¹²El CGIAR adapta el concepto de «rigor de investigación» del Instrumento de Evaluación RQ+ del International Development Research Centre (IDRC, 2022) a las necesidades del CGIAR, enmarcándolo de manera que puede tenerse en cuenta incluso en la ausencia de productos específicos.

3.1

Cuatro dimensiones para evaluar la QoS y mapear otros criterios de evaluación continuación

La Tabla 3 mapea las cuatro dimensiones de la QoS a los siete criterios de evaluación: demuestra cómo «encajan» las cuatro dimensiones con siete criterios de evaluación, que enmarcan cómo se hacen los juicios evaluativos para responder a las preguntas evaluativas centrales en las evaluaciones de proyectos y desempeño de la cartera del CGIAR. Esto significa que las dimensiones sugieren preguntas de evaluación que pueden implementarse junto con múltiples criterios, por lo que existen muchas rutas para la exploración evaluativa de la ciencia. Las preguntas sugeridas por los criterios de evaluación se presentan en mayor detalle en el [Paso 2](#) y el [Anexo 3](#).

Table 3: Ejemplo de mapeo de siete criterios de evaluación en cuatro dimensiones de la QoS

QoS			CRITERIOS DE EVALUACIÓN					
			Relevancia	Coherencia	Eficiencia	Eficacia	Sostenibilidad	Impacto
DIMENSIONES	Diseño	X	X	X				X
	Insumos	X	X	X				
	Proceso	X		X	X	X	X	X
	Resultados	X	X			X	X	X

En todas las dimensiones de la QoS, se requiere prestar atención a los estándares y principios de evaluación (Anexo 5). Por ejemplo:

- La consideración del principio de género, diversidad e inclusión (GDI) sería adecuado en diseño, insumos y procesos (2022, EF). El diseño y la realización de la evaluación, la creación de equipos y la presentación de informes se esfuerzan por abordar los parámetros de GDI de manera exhaustiva. Las evaluaciones considerarán quienes son las personas involucradas y los beneficiarios del trabajo. La evaluación consideraría la composición de un equipo de investigación en términos de diversidad de género, nacionalidad, edad y disciplina, así como la participación de mujeres, investigadores de los LMIC y comunidades locales en el diseño y la ejecución de la investigación.

- Los principios de legitimidad, participación, y transparencia abordan el papel cada vez más crítico de las alianzas para facilitar un buen diseño, procesos y resultados de la investigación. Las evaluaciones cuestionan las percepciones y prácticas, al invitar a informantes relevantes y utilizar procesos consultivos durante el diseño de preguntas de evaluación y la selección de métodos. También aseguran el proceso de entrega de ciencia al incluir a representantes de los usuarios finales e intermedios de los productos, cuando es apropiado y factible.

Los informes de evaluación inicial y final deberían incluir cómo se abordan los estándares y principios y aportar consideraciones respecto de mitigaciones y limitaciones potenciales y reales.

3.2 Preguntas clave de evaluación de la QoS

Se recomiendan las siguientes tres preguntas clave de evaluación (EQ), especialmente cuando se usa un criterio designado de evaluación de la QoS (Figura 4 y Tabla 3).

Figura 4: Tres preguntas de evaluación para poner en práctica los criterios de evaluación de la QoS



La EQ.1 captura el diseño, cómo de apropiado y receptivo es para abordar los desafíos del desarrollo. La evaluación y el juicio se realizan en términos de estándares comúnmente aceptados en un campo temático designado, la idoneidad técnica de la agenda y la estrategia de investigación. Esta dimensión se solapa con los criterios de evaluación de relevancia y coherencia hacia el impacto y aborda el elemento de credibilidad (consulte la Tabla 2). El diseño evaluativo debe describir cómo se entrelazarán los criterios de la QoS y la relevancia. Esta pregunta proporciona al equipo de evaluación un espacio para analizar en profundidad la contextualización del esfuerzo científico.

La EQ.2 considera la calidad de los insumos y procesos. Solo tiene sentido si se compara con puntos de referencia claros de la industria y/o estándares u objetivos de calidad del CGIAR (p. ej., diversidad de género entre el personal científico, requisitos para infraestructura y

materiales). La credibilidad resulta de la idoneidad de los insumos; el proceso confiere legitimidad. La EQ.2 en gran parte se solapa con el criterio de evaluación de la eficacia, en relación al elemento de legitimidad.

La EQ.3 evalúa la contribución de los resultados a un campo de investigación, dominio, disciplina o agrupación transdisciplinaria de estos. Responder a la EQ.3 requiere una comprensión del alcance del esfuerzo científico dados los enfoques interdisciplinarios y transdisciplinarios que se utilizan a menudo para cumplir la misión del CGIAR. Los equipos de evaluación y sus diseños deben abordar la EQ.3 en base al entendimiento que transformar el sistema requiere formas transdisciplinarias de trabajo porque muchas disciplinas distintas pueden estar involucradas en una intervención generadora de resultados. Esto puede solaparse con los criterios de evaluación de la eficiencia y la coherencia.

3.3 Calendarización de la evaluación de la QoS en el CGIAR

Se ha demostrado que los productos seleccionados y la mayoría de los resultados sociales de la investigación suelen tardar cinco, diez o incluso veinte años en materializarse por completo (Langfeldt, 2015). Este es el caso, en particular, de la investigación relacionada con la tecnología, es decir, el fitomejoramiento. Sin embargo, parece cada vez más que la investigación de tipo de sistemas está orientada a evaluar y lograr impactos en tiempo real. La calendarización para evaluar la QoS en el contexto de la evaluación de procesos y desempeño dependería del tipo de intervención y de las dimensiones a ser evaluadas, así como del ciclo general del proyecto. Pueden

calcularse diferentes métodos e indicadores en diferentes momentos durante o más allá de la duración de una intervención. Con frecuencia, p. ej., es útil evaluar insumos y procesos antes de la siguiente fase de una intervención, para respaldar la gestión adaptativa y la corrección del rumbo. Por tanto, una evaluación integral de la QoS, que incorpore las cuatro dimensiones, es mejor al final de un ciclo de proyecto de una intervención de tipo AR4D.

La Figura 5 muestra un ejemplo de intervención de CGIAR con ciclos de financiación de tres años diseñados para generar resultados a lo largo de tres fases consecutivas (como es el caso de las iniciativas de financiación conjunta de CGIAR).

Figura 5: Cronograma típico para la duración de un proyecto R4D*

Fases	Año 0	Año 1	Año 2	Año 3	Año 4
Diseño de intervención					
Implementación					
Monitoreo					
Evaluación de procesos/desempeño					

*Publicación de artículos relacionados y ventanas de citación con factor de impacto, aplicadas a 44 iniciativas y programas del CGIAR (2022).

Se recomienda un enfoque triple para evaluar la oportunidad y solidez de la evaluación con las siguientes consideraciones:

- **Año 2:** El análisis de evaluabilidad o evaluación intermedia equilibraría los métodos cualitativos y cuantitativos para evaluar las dimensiones de insumos y procesos. Cuando los recursos lo permiten, una evaluación intermedia identificaría las prácticas susceptibles de mejora, lo que permitiría la gestión adaptativa de insumos y prácticas en procesos en curso. Gran parte de los indicadores cuantitativos de resultados solo se pueden calcular tras un período temporal más largo (año 3-8). Sin embargo, un indicador como el índice H podría usarse ya en esta etapa como una medida de la productividad de los científicos involucrados en la intervención.
- **Año 3-5:** La evaluación final (procesos/desempeño) también priorizaría los métodos cualitativos para evaluar las dimensiones de insumos y procesos y los indicadores cuantitativos, p. ej., un conjunto restringido de indicadores bibliométricos.

- **Año 8:** La evaluación integral y específica de métodos mixtos de una cartera de proyectos, para captar los resultados sociales logrados a medio plazo (p. ej., adopción de publicaciones en documentos relacionados con políticas, bienes públicos internacionales) o el impacto acumulado de citas de artículos de investigación transformadores. La Figura 6 ofrece sugerencias sobre la evaluación de procesos y desempeño de las dimensiones de la QoS por criterios de evaluación, dependiendo del ciclo del proyecto y la duración general de los proyectos de R4D. Sin embargo, cambios en los cronogramas sugeridos son posibles si una intervención específica permite una evaluación temprana de las dimensiones seleccionadas, es decir, el diseño a medio plazo. En particular, los recuadros en azul hacen referencia a la evaluación de la QoS. El azul claro incluye los aspectos que pueden evaluarse antes del final de la intervención, mientras otros deben evaluarse tras el final de la intervención.

Una evaluación integral de la QoS, que incorpore las cuatro dimensiones, es mejor al final de un ciclo de proyecto de una intervención de tipo AR4D.

Figure 6: Oportunidad para la evaluación de cada criterio

Criterios de evaluación/Ciclo del proyecto	Año 0	Año 1	Año 2	Año 3	Año 4
Relevancia					
Coherencia					
Eficacia					
Eficiencia					
Sostenibilidad					
Impacto					
Calidad de la ciencia					

3.4 Garantía de calidad

Para garantizar la solidez y la credibilidad de una evaluación, integrada en el Marco de evaluación del CGIAR y alineada con el proceso de evaluación estándar, debe seguirse en todo momento un sistema multicapa de garantía de calidad (QA).

La triangulación es un método utilizado para aumentar la credibilidad y validez de los hallazgos de la evaluación y es esencial en el diseño, la implementación y la QA de la evaluación. Diferentes tipos de triangulación son apropiados en diferentes pasos de evaluación de la QoS.

(1) triangulación de investigadores (miembros del equipo de evaluación), definida como el uso de diferentes expertos en la recopilación de datos y el proceso de análisis.

(2) triangulación de teorías, lo que implica el uso de múltiples enfoques durante la evaluación.

(3) triangulación metodológica, es decir, el uso de más de un método para la recopilación de datos, cuando sea posible.

(4) triangulación de datos, es decir, el uso de diferentes fuentes de datos para reforzar el resultado.

Cuadro de texto 2: Integración de la garantía de calidad en el proceso de evaluación

Para tener solidez y credibilidad en una evaluación, se recomienda un equipo diversificado de expertos multidisciplinarios. Cada equipo debe estar compuesto por un líder de equipo (preferiblemente, un evaluador profesional) y expertos en la materia (SME). Estos últimos son líderes de opinión en ámbitos relevantes para la Estrategia de investigación e innovación de CGIAR 2030. Este equipo, complementado con una rigurosa revisión por pares y otros procedimientos de control de calidad, garantiza la credibilidad del proceso evaluativo y los resultados (Sección 4).

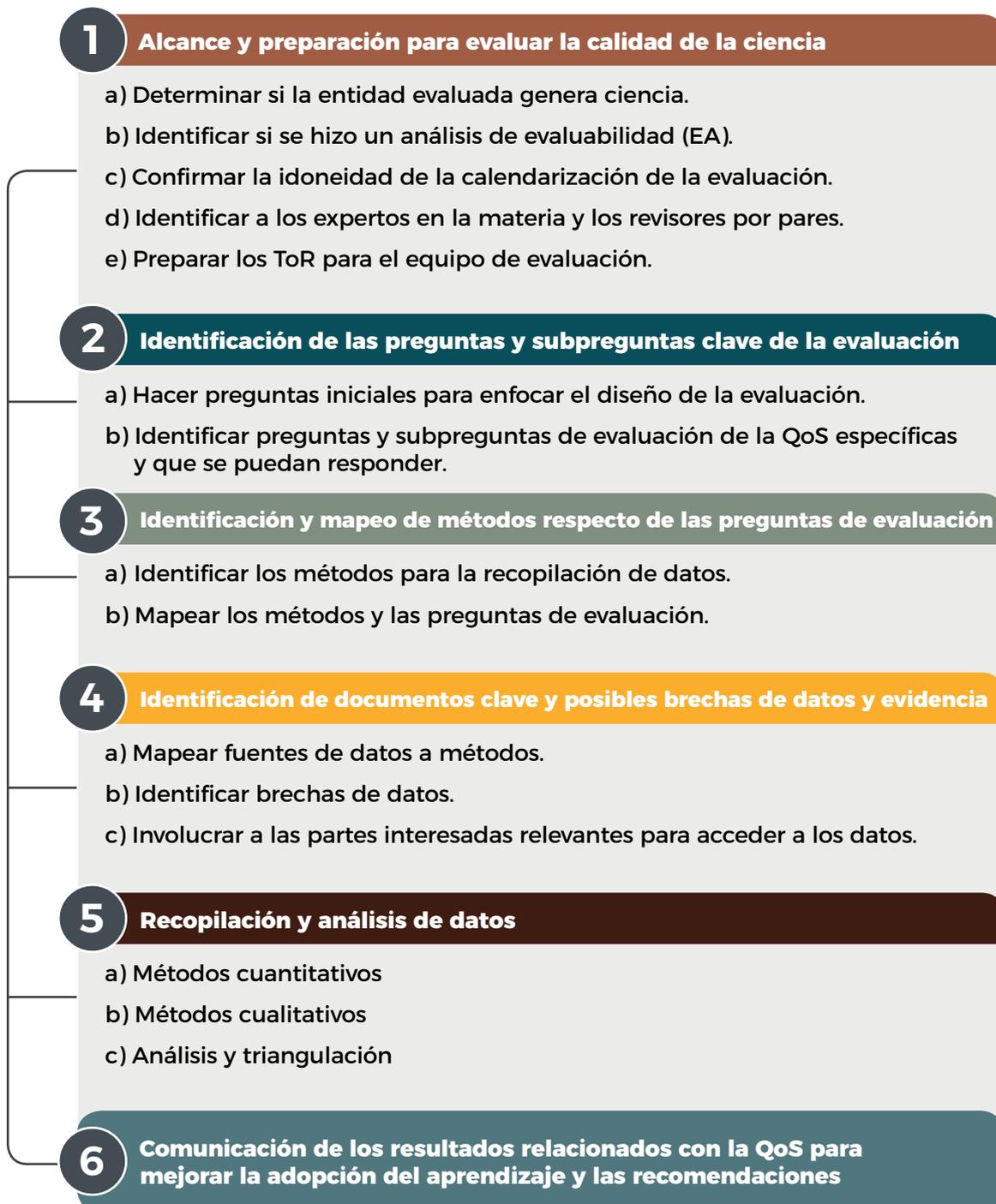
Además de un equipo diversificado que incluya a los SME en el equipo de evaluación, también se deben incluir a revisores externos expertos en el tema a evaluar. Participarán en al menos dos pasos del proceso de evaluación. Complementando la revisión por parte de IAES/Evaluación, los revisores por pares de la evaluación verifican la elección de metodología en cuanto a calidad y solidez técnica y revisan los informes de evaluación iniciales y los borradores. También revisan el módulo o componente u otros entregables intermedios clave. Los problemas que surjan durante los pasos de QA deben ser mitigados en la medida de lo posible por los SME y el líder del equipo de evaluación con documentación clara (Paso 1)

En la Sección 4, se describe un proceso de seis pasos clave para facilitar la evaluación de la QoS, ya sea como parte de una evaluación más amplia o muy dirigida al criterio de evaluación de la QoS como tal. Los Pasos 3 y 4 abarcan un menú de métodos y el Paso 5 se centra el papel de la triangulación de datos.

Para cualquier evaluación dada, un análisis de la evaluabilidad, u otra etapa previa al alcance o la planificación recomendaría una elección entre los siete criterios del CGIAR que sean más apropiados para el propósito específico de la evaluación. Se puede implementar el uso de un criterio de evaluación de la QoS designado (preferiblemente junto con criterios de evaluación de relevancia y eficacia como mínimo), o la QoS se puede descubrir de indirectamente a través de combinaciones de los otros seis criterios de evaluación (consulte la

Tabla 3). La ruta a seguir se determina según la calendarización y el propósito de la evaluación. En cualquier caso, se sugiere un enfoque de seis pasos para evaluar la QoS. Cada paso (Figura 7) presenta un punto de decisión para asegurar que continúe la aplicabilidad de las dimensiones relevantes (diseño, insumos, procesos y resultados- Sección 3.1) y los métodos elegidos. Para evaluaciones guiadas por múltiples criterios, estos pasos se implementarían junto con todos los criterios de evaluación en uso.

Figura 7: Seis pasos clave para la evaluación de la QoS



Paso 1: alcance y preparación para evaluar la QoS

Este primer paso determina si la entidad evaluada genera o no investigación científica. Determinadas intervenciones evaluadas, p. ej., plataformas, no están obligadas per se a generar ciencia. Por tanto, los objetivos de la entidad evaluada impulsan la evaluación; y, la pregunta rectora central es «¿El objetivo de la entidad evaluada era generar ciencia?»¹³ En caso afirmativo, la QoS es un determinante importante de su eficacia potencial ([Anexo 2](#)). Así, es necesario el Paso Uno para comprender si el criterio de la QoS debe aplicarse respecto de la intervención sujeta a evaluación. En caso afirmativo, implica subpasos preparatorios subsecuentes.

1a. Determinar si la entidad evaluada genera ciencia.

- Haga la pregunta: «¿El objetivo de la entidad evaluada era generar productos y resultados científicos para el desarrollo?». En caso negativo, no incluya el criterio de la QoS y siga otra ruta hacia el diseño de la evaluación. En caso afirmativo, entonces...
- ... identifique la etapa del ciclo del proyecto en la que se encuentra la entidad evaluada, es decir, entrega de productos, para enfocarse en las dimensiones: diseño, insumos, procesos y resultados (Sección 3) para controlar el alcance, diseñar las preguntas de evaluación y seleccionar los métodos apropiados.

Ejemplo: Podría hacerse una clara distinción en la **Cartera de investigación del CGIAR** entre los programas que generan ciencia ([Revisiones evaluativas de 2020 de los programas de investigación del CGIAR](#)) y aquellos que coordinan la entrega de ciencia (p. ej., las plataformas CGIAR sobre [Big data en agricultura](#) y [Excelencia en fitomejoramiento](#), cuyos objetivos se centraban en el apoyo y la coordinación en torno a los resultados y procesos científicos. Específicamente, *big data* en Agricultura tenía como objetivo aprovechar las capacidades de *big data* para acelerar y mejorar el impacto de la investigación agrícola internacional, pero no generó ciencia directamente, mientras que Excelencia en fitomejoramiento, con un fuerte enfoque en los datos, se centra en impulsar el cambio en lugar de en la entrega de ciencia.

1b. Identificar si se realizó o no un análisis de evaluabilidad (EA). Una EA identifica si una entidad a ser evaluada está preparada para ello mediante la valoración de si una EA proporcionó y/o capturó la siguiente información:

- La TdC (bien definida) asociada a cada intervención y los objetivos originales.
- Las necesidades, políticas y prioridades de los usuarios, incluidos los socios e instituciones mundiales, regionales y nacionales.
- La importancia, el significado y la utilidad de la ciencia generada en un contexto específico.

En caso de no haberse realizado un EA, estas áreas deben abordarse antes de llevar a cabo la evaluación (como un proceso separado) o abordarse en el diseño de la evaluación.

1c. Confirmar la idoneidad de la calendarización de la evaluación. Para orientaciones específicas, consulte la [Sección 3](#). Por ejemplo, podría ser útil para los usuarios de la evaluación aprender sobre la idoneidad de los insumos y procesos antes del final de la intervención o el comienzo de la siguiente fase.

1d. Identificar a los expertos en la materia y los revisores por pares. En CGIAR y contextos similares de AR4D, se deben incluir expertos calificados que tengan conocimientos sobre investigación para el desarrollo en CGIAR y en contextos de AR4D más amplios.

Además:

- Revisores por pares: revisores por pares en un Grupo de referencia de evaluación (ERG para el IAES) que son expertos en evaluación para revisar el diseño de la evaluación y los resultados relevantes.
- Expertos en la materia (los SME): científicos con experiencia que es específicamente relevante para la ciencia que se está evaluando.

1e. Preparar los ToR para los miembros del equipo de evaluación. Los ToR para el líder del equipo de evaluación y los SME deben exigir explícitamente que estén familiarizados con estas pautas sobre la aplicación del Marco de QoR4D a las evaluaciones de procesos y desempeño.

¹³En el caso de CGIAR, en coherencia con la [Estrategia de investigación del CGIAR 2022-2030](#) «...para entregar ciencia e innovación que avancen en la transformación de los sistemas alimentarios, terrestres e hídricos en una crisis climática... Innovación basada en la ciencia: el desarrollo compartido de conjuntos de productos de conocimiento, tecnologías, servicios y otras soluciones a lo largo de un camino de escalamiento. El CGIAR trabajará con socios en innovaciones que incluyen genética, prácticas de gestión agroalimentaria, ciencias sociales y soluciones institucionales, ciencias biofísicas y soluciones, bases de datos y herramientas. Las actividades incluirán diseño participativo, pruebas y pilotaje, trabajando en estrecha colaboración con socios del sector privado y organismos reguladores, avanzando el entorno propicio y proporcionando una arquitectura global para la investigación agrícola internacional colaborativa».

Paso 2: Identificación de las preguntas y subpreguntas clave de la evaluación

Deben prepararse dos conjuntos de preguntas: (1) preguntas que informan el diseño de la evaluación y (2) aquellas que abordan la intención evaluativa (es decir, preguntas de evaluación).

2a. Hacer preguntas iniciales para enfocar el diseño de la evaluación. El primer conjunto de preguntas se enfoca, en general, en el diseño de la evaluación al aclarar qué debe evaluarse. Para esto, utilice la TdC definida y solicite al representante del comisario y al gerente del programa que la elaboren preguntando:

- ¿Cómo contribuye un resultado a avanzar un campo de investigación, dominio, disciplina o agrupación transdisciplinaria de estos?

- ¿La ciencia generada es relevante para los desafíos de desarrollo identificados?
- ¿Cómo se mapean las innovaciones (de tres tipos) a los procesos y resultados? ([Tabla 1](#))

2b. Identificar preguntas y subpreguntas de evaluación de la QoS que son específicas y pueden ser contestadas a través de un ejercicio evaluativo (Tabla 4). En base a la información obtenida durante los pasos anteriores, desarrolle preguntas y subpreguntas clave de evaluación relacionadas con la QoS, utilizando el [Anexo 5](#) y el [Anexo 6](#)¹⁴.

Para guiar aún más la creación de subpreguntas y parámetros de evaluación en las dimensiones de la QoS, la Tabla 7 presenta un menú de temas por estas dimensiones, con indicadores y criterios de evaluación sugeridos.

Table 4: Criterios de evaluación con ejemplos de preguntas de evaluación

Criterios de evaluación	Ejemplo de pregunta de evaluación	Dimensión de la QoS abarcada (diseño, insumos, proceso, resultados)
QoS	EQ1: ¿El diseño de la investigación es adecuado para los desafíos de desarrollo en el contexto? EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima? EQ3: ¿Cómo contribuyen los resultados de la intervención al avance de la ciencia? (en línea con la concepción íntegra de relevancia según la QR4D)	TODAS
Relevancia	- ¿Existe evidencia de demanda (continua) del programa por parte de los beneficiarios previstos? - ¿Es el programa coherente con el PRMF y el Marco de resultados, y la agenda de reforma acordada del CGIAR?	Diseño, Resultados
Coherencia	¿El diseño de la intervención es coherente con otras intervenciones en la cartera de investigación/el país/el sector?	Diseño, Insumos
Eficiencia	- ¿La financiación era adecuada y oportuna? - ¿Las instalaciones/los servicios son adecuados, se utilizan adecuadamente? - ¿La composición de los equipos de investigación era suficientemente diversa (inclusiva en términos de género, edad/jóvenes investigadores y nacionalidad)?	Insumos, Proceso
Eficacia	- ¿La intervención logró sus objetivos y resultados? ¿Los entregables están posicionados para ser aceptados? - ¿Los roles y las responsabilidades se definieron claramente y se implementaron según lo planeado, en línea con las esferas de control e influencia de las TdC? - ¿Alguna actividad en la cartera actual debe modificarse, suspenderse o incluirse para mejorar la eficacia probable del programa?	Proceso, Resultados
Sostenibilidad	- ¿La implementación de las TdC del programa y los supuestos que subyacen a estas teorías incluyen aspectos de sostenibilidad? - ¿La contribución generada por la intervención es escalable y es probable que continúe? - ¿Se han tenido en cuenta las compensaciones entre los diferentes resultados a largo plazo en el diseño y la implementación del programa, p. ej., respecto de la sostenibilidad ambiental?	Resultados
Impacto	Cuál fue el impacto de los estudios generados/las intervenciones? ¿Tuvo un efecto transformador?	Resultados

¹⁴The data collection matrix covers dimensions of design, input, process and outputs by specific evaluation criteria.

Paso 3: Identificación y mapeo de métodos respecto de las preguntas de evaluación

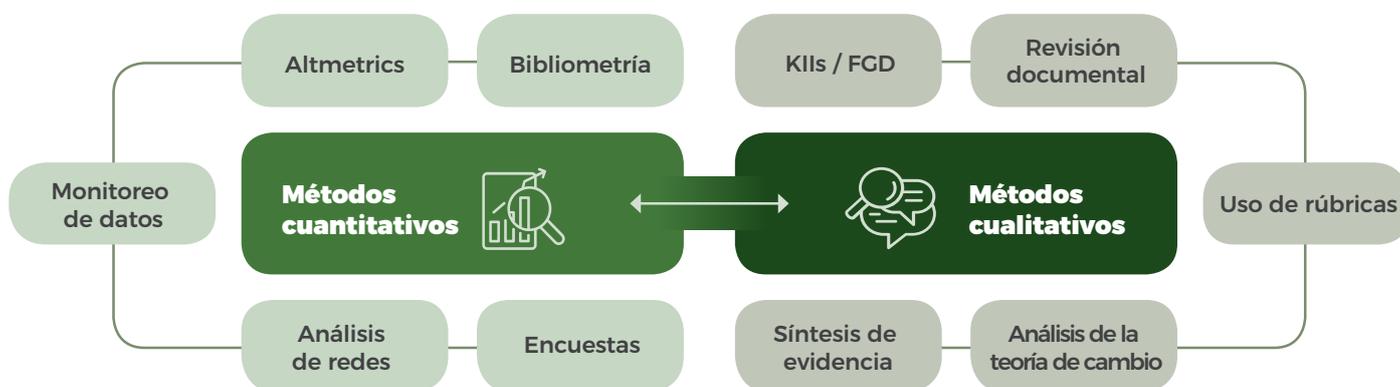
El Marco de evaluación, coherente con los estándares de evaluación, aconseja el uso de métodos mixtos en las evaluaciones de procesos y desempeño, lo que también incluye evaluar la QoS. Asimismo, se deben considerar fuentes de datos primarias y secundarias.

3a. Identificar los métodos para la recopilación de datos

Los métodos deben seleccionarse en función de las preguntas de evaluación, la disponibilidad de datos, la preferencia de los usuarios clave (es decir, qué tipo de datos consideran creíbles los usuarios principales), y la

calendarización de la evaluación. Por ejemplo, las preguntas que se centran en el proceso, la capacitación o la comunicación pueden basarse en datos cuantitativos (p. ej., número de personas capacitadas, número de métodos o herramientas, puntuaciones en las pruebas de comprensión de los alumnos) y/o cualitativos (p. ej., calidad de la capacitación, utilidad para las actividades planificadas, relevancia para el público objetivo). La Figura 8 ofrece un vistazo al menú de los métodos disponibles para la recopilación y el análisis de datos.

Figura 8: Métodos de evaluación para la recopilación y el análisis de datos



3b. Mapear los métodos y las preguntas de evaluación

Los miembros del equipo de evaluación, especialmente los SME, deben participar en el mapeo de las preguntas y subpreguntas clave de la evaluación a los métodos de evaluación para la recopilación y el análisis de datos (Figura 8, Tabla 5, Cuadro de texto 3 y Anexo 3). La matriz

de diseño de evaluación, incluido el criterio de la QoS (Anexo 6), ha sido desarrollada para identificar y agrupar las fuentes de información a métodos cuantitativos y cualitativos. El líder del equipo de evaluación involucra al equipo a través del desarrollo de la matriz de diseño de la evaluación, un elemento clave del informe inicial de la evaluación.¹⁶

Tabla 5: Ejemplos de preguntas y subpreguntas de evaluación por dimensiones de la QoS, fuentes de datos y métodos*

Dimensiones de la QoS	Pregunta de evaluación	Ejemplos de subpreguntas de evaluación	Fuentes de datos, métodos
Diseño	¿El diseño de la investigación es apropiado y está articulado con claridad?	<ul style="list-style-type: none"> - ¿Son las preguntas y la metodología de investigación adecuadas para el propósito y alineadas con el problema de investigación? - ¿Cómo de interconectados están el diseño de la investigación y los ODS? - ¿Existe evidencia de cómo y qué socios participaron en el codiseño? - ¿Hay un claro vínculo entre el plan MEA/El y la TdC en el diseño de la iniciativa de investigación? 	Propuestas de iniciativas e informes; informes de ISDC de revisión de la iniciativa; TdC (original y revisiones); revisiones de propuestas ex ante de ISDC <i>Primarias:</i> Entrevistas
Insumos	¿Hasta qué punto los insumos necesarios eran adecuados y suficientes para entregar los productos y resultados planificados?	<ul style="list-style-type: none"> - ¿La composición de los equipos de ejecución de la investigación era adecuadamente diversa? - ¿Las infraestructuras físicas de investigación (p. ej., laboratorios, parcelas experimentales, etc.) eran adecuadas? - ¿La capacitación del equipo de investigación y los socios abordó las necesidades respecto del trabajo planificado, incluidos los aspectos no científicos? 	Informes de iniciativas; SNA; revisiones de presupuesto contra planes; evaluaciones de necesidades y registros de capacitación <i>Primarias:</i> Visitas de campo/laboratorio de expertos; entrevistas
Procesos	¿En qué medida el proceso de gestión aseguró la QoS, incluida la credibilidad científica y la legitimidad de la investigación y las operaciones?	<ul style="list-style-type: none"> - ¿Cuáles eran los niveles de confianza, com promiso e involucración de los diferentes socios? - ¿Se siguieron las políticas sobre la tutoría y formación del personal de investigación junior? - ¿Se usaron los datos de desempeño y monitoreo para la gestión adaptativa? 	Informes de iniciativas; registros de reuniones; políticas internas <i>Primarias:</i> Entrevistas, FGD

*Consultar el [Anexo 3](#) para más detalles

¹⁶La guía del Informe inicial (IR) para CGIAR está en proceso de finalización: consulte el ejemplo de IR de la [Evaluation of CGIAR Platform for Big Data in Agriculture: Inception Report \[Evaluación de la Plataforma de big data en agricultura del CGIAR: informe inicial\]](#) y el blog relacionado en [Evaluation Inception Reports - The CAS Approach \[Informes iniciales de evaluación: el enfoque de CAS\]](#)

Tabla 5: Ejemplos de preguntas y subpreguntas de evaluación por dimensiones de la QoS, fuentes de datos y métodos (continuación)

Dimensiones de la QoS	Pregunta de evaluación	Ejemplos de subpreguntas de evaluación	Fuentes de datos, métodos
Resultados	¿Cuál es la calidad de los resultados de la investigación, tales como variedades mejoradas, herramientas de conocimiento y publicaciones de alta calidad? ¿Cómo contribuyen los resultados de las intervenciones al avance de la ciencia?	- ¿Cuál es la contribución de los resultados a las innovaciones basadas en la ciencia, la capacitación específica y el asesoramiento sobre políticas? - ¿Se comunicaron/difundieron claramente los hallazgos de la investigación y los resultados relacionados? - ¿Cómo se reflejan el GDI o las preocupaciones ambientales en los resultados? ¿Pueden estos productos tener aplicabilidad más amplia y potencial de impacto a escala?	Bibliometría; Altmetrics; informes de iniciativas; TdC; análisis y evaluación de expertos; estadísticas de descargas; tendencias en las redes sociales <i>Primarias:</i> Entrevistas, FGD

Los métodos a tener en cuenta en la evaluación de la calidad de la ciencia se desglosan por tipo en el Cuadro de texto 3, con detalles adicionales sobre las fortalezas y limitaciones de cada método en el Paso 5.

Cuadro de texto 3: Ejemplos de métodos cuantitativos y cualitativos recomendados para evaluar la calidad de la ciencia

Métodos cuantitativos podrían incluir **encuestas, análisis bibliométrico, Altmetrics, SNA y datos de monitoreo**. El análisis cuantitativo mitiga los juicios subjetivos de los SME. El proceso de hacer ciencia es difícil de medir a través de métodos cuantitativos, a pesar del creciente número de indicadores bibliométricos, y depende en gran medida de métodos cualitativos. (Consulte detalles adicionales en el [Paso 5](#))

Métodos cualitativos podrían incluir **revisiones de expertos, entrevistas, grupos focales FGD, análisis de la TdC, rúbricas y síntesis de evidencia**. El uso de métodos cualitativos es clave para evaluar la ética y la inclusividad de un proceso, su integración del aprendizaje, y hasta qué punto la QoS radica en el contenido generado y se muestra en los resultados. Algunos métodos cualitativos pueden ser más propensos al sesgo. (Consulte detalles adicionales en el [Paso 6](#))

Paso 4: Identificación de documentos clave y posibles brechas de datos y evidencia

Determinación de qué recopilación y análisis de datos primarios y secundarios se debe realizar en este paso y, según las fuentes de datos disponibles y los métodos preferidos, debe determinarse cualquier brecha en las fuentes de datos y la viabilidad de los métodos. Si es

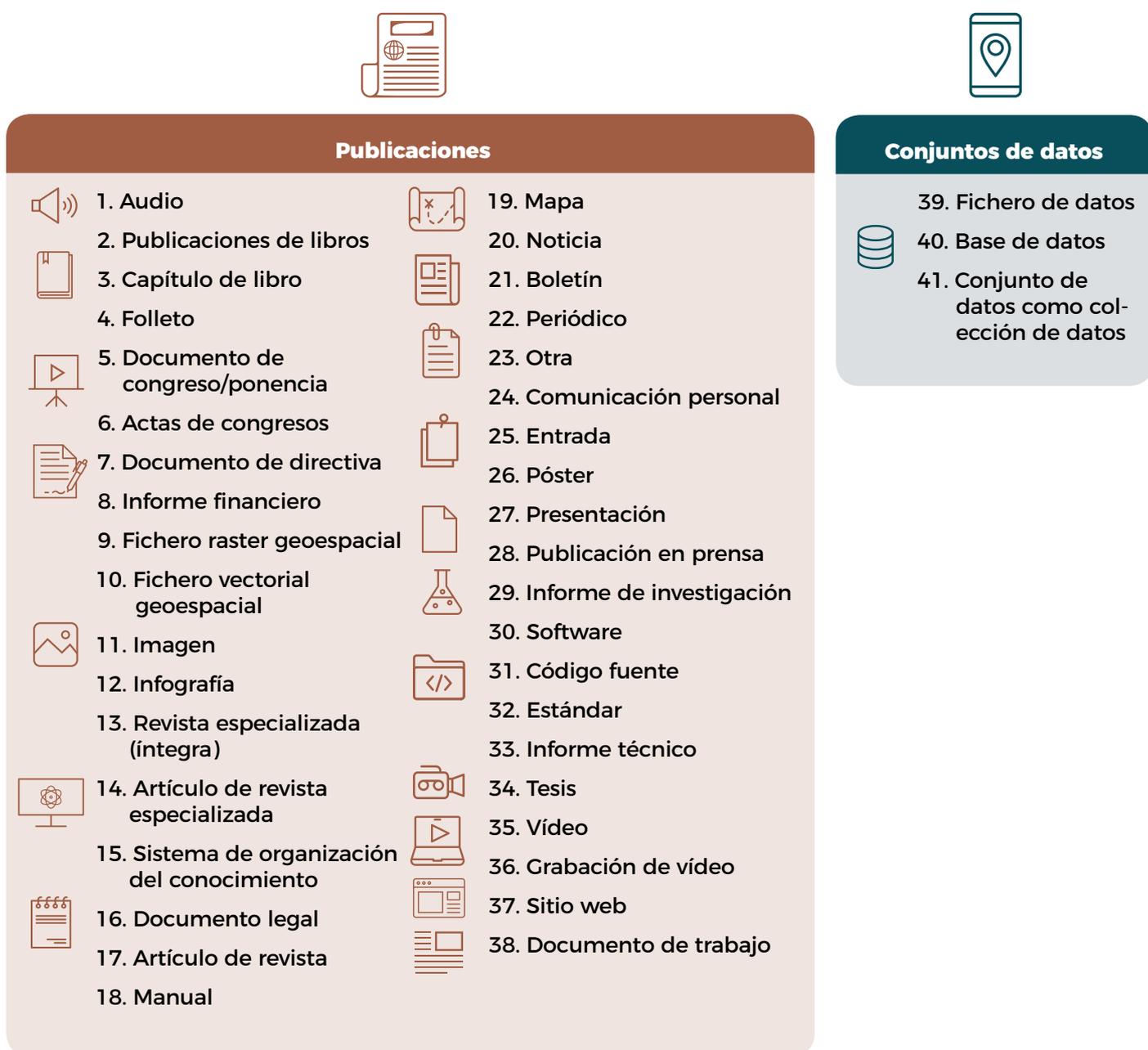
Cuadro de texto 4: Lista de verificación para la recopilación de documentos e información necesario, el equipo de evaluación debe colaborar con las partes interesadas relevantes para el acceso a datos y evidencia. En este paso, la lista de verificación del Cuadro de texto 4 complementará la información recopilada durante la determinación del alcance, si no durante la EA. Por último, un resultado podría ser la creación de nuevas instituciones (Glosario de PRMF/MEL/IA 2021; IAES).

Cuadro de texto 4: Lista de verificación para la recopilación de documentos e información

- ✓ **Recopilar los nombres y la información de contacto de las principales partes interesadas y los SME (para KII/FGD)**
- ✓ **Revisar la disponibilidad de la QoS y la información relacionada con el proyecto si no se hizo anteriormente en la EA**
- ✓ **Dependiendo de la disponibilidad y calidad de la documentación, determinar si usar fuentes de datos externas y subcontratar algunos elementos de la evaluación de la QoS, p. ej., el análisis bibliométrico (consulte el [Anexo 6](#))**

Como ejemplo, para facilitar el **análisis bibliométrico/Altmetrics**, la Figura 9 presenta los resultados recomendados por el [CG Core](#). El CG Core tiene como objetivo describir todo tipo de productos de información publicados por los diferentes centros del CGIAR. Un ejemplo de un indicador relacionado con los resultados en los activos de datos: los tipos de activos de datos presentados son recomendados por el esquema de metadatos, alineado con los estándares de la industria. Si bien la lista contiene resultados comunes, también se puede utilizar para generar con la entidad evaluada ideas respecto de posibles fuentes de datos.

Figura 9: Lista de activos de datos (resultados) en el CG Core agrupados en dos categorías de alto nivel



Paso 5: Recopilación y análisis de datos

Con la información disponible sobre la evidencia y las brechas, y las necesidades conocidas para la recopilación de datos primarios o secundarios, el siguiente paso implica considerar los métodos de recopilación y análisis de datos específicos. Sin embargo, antes de recopilar y analizar datos, se debe tener en cuenta las fortalezas y limitaciones de los diferentes métodos de recopilación de datos, descritos aquí.

5a. Uso de métodos cuantitativos

La Tabla 6 establece las fortalezas y limitaciones respectivas de varios métodos cuantitativos a considerar para la evaluación de la QoS.

Tabla 6: Menú de métodos cuantitativos sugeridos para la evaluación de las dimensiones de la QoS, con fortalezas y limitaciones

Métodos ¹⁷	Definición general	Fortalezas	Limitaciones
Análisis bibliométrico¹⁸	Utilizado tradicionalmente para evaluar los resultados y su eficacia; es un análisis estadístico de libros, artículos y otras publicaciones, en particular aquellas con contenido científico.	Incluye algunos de los indicadores de impacto científico más ampliamente aceptados, p. ej., impacto dentro de un campo de investigación (p. ej., factor de impacto, índice de citas); proporciona una buena indicación de la QoS, ya que los artículos publicados han superado un umbral de alta calidad tras su revisión por pares; su valor es reconocido por financiadores. En la actualidad ofrece indicadores para una amplia gama de dimensiones que permiten también la evaluación de procesos e insumos. Los nuevos indicadores bibliométricos permiten medir la interdisciplinariedad, la equidad de género, la preimpresión como práctica de ciencia abierta o la prevalencia de colaboraciones multinacionales complejas.	Limitarse a artículos revisados por pares medidos a través de indicadores bibliométricos no tendría en cuenta la ciencia bien realizada que no genera resultados significativos y que, por tanto, es difícil de publicar. No todos los productos de ciencia, innovación e investigación están incluidos en las bases de datos bibliográficas; las bases de datos bibliográficas provenientes de diferentes fuentes pueden no estar armonizadas; puede haber una atención exagerada a un autor específico; períodos prolongados (los resultados de las inversiones en investigación agrícola pueden tardar décadas en hacerse visibles); proporciona poca información sobre el alcance de las políticas, la relevancia contextual, la sostenibilidad, la innovación y la escalabilidad de la contribución; la dependencia exclusiva de la bibliometría podría omitir una visión más completa del contexto; las medidas bibliométricas pueden estar sesgadas dependiendo del dominio de investigación (es decir, puede ser inapropiada la comparación entre dominios).
Altmetrics	Altmetrics se utiliza para monitorear el alcance e impacto de los estudios y la investigación a través de interacciones en línea.	Altmetrics son datos cuantitativos que complementan las métricas tradicionales, basadas en citas. Más rápidas de acumular que las métricas basadas en citas; pueden capturar impactos más diversos que las métricas basadas en citas; son aplicables más allá de artículos de revistas y libros.	Altmetrics no se puede usar individualmente porque no cuenta la historia completa; la comunidad científica no la reconoce de forma generalizada; existe la posibilidad de jugar con ella de forma que podría sesgar los resultados; es una herramienta relativamente nueva, y son necesarios más estudios para sacarle mejor provecho.
Análisis de redes sociales (SNA)	Proceso de investigación de estructuras sociales usando análisis de redes y teoría de grafos.	Muy práctico y visualmente atractivo; se puede hallar mucha información al interpretar gráficos de red; también muestra brechas en las conexiones.	La pregunta de investigación debe ser clara para saber exactamente qué se debe mapear. La interpretación de la red puede ser compleja. Puede ser necesario confiar en expertos externos para los métodos y hacer la interpretación de datos a nivel interno para tener en cuenta el contexto del CGIAR.
Encuestas	Recopilación de datos de un grupo predefinido de encuestados.	Es relativamente fácil de crear y no necesariamente requiere expertos externos; existen muchas herramientas de bajo costo utilizadas para diseñar encuestas; se puede llegar a un gran número de encuestados.	La tasa de respuesta puede ser limitada y los encuestados pueden tener sesgos; debe compararse a través de los canales correctos para limitar los sesgos; la fatiga de la encuesta podría llevar a un sesgo de respuesta; los encuestados pueden interpretar las preguntas de manera diferente si su redacción no es clara.
Datos de monitoreo del rendimiento	El PRMS ¹⁹ abarca la planificación, el monitoreo y la presentación de informes dentro de CGIAR.	Elementos PRMS - Cuadro de indicadores de resultados, entre otros, convierten información relacionada en productos de evaluación. Ver https://www.cgiar.org/dashboards/	La calidad de los datos y la información en los cuadros de indicadores depende de los mecanismos de garantía de calidad, que variarán según el cuadro de indicadores. La calidad variable de los datos se evaluó durante las Revisiones de los CRP 2020.

¹⁷Consulte detalles adicionales sobre los métodos seleccionados en el [Anexo 3](#)

¹⁸<https://cas.cgiar.org/evaluation/publications/bibliometric-analysis-evaluate-quality-science-context-one-cgiar>

¹⁹Este listado de indicadores se basó en el [RQ+ Assessment Instrument \(IDRC, 2022\)](#) e incluye indicadores utilizados previamente para las Revisiones de los CRP 2020.

5b. Uso de métodos cualitativos en ciencias sociales recopila y trabaja con datos no numéricos y busca interpretar el significado de estos datos que ayudan a entender la vida social a través del estudio de poblaciones o lugares específicos, incluso a través de palabras e imágenes, en contraste con los métodos cuantitativos centrados en cifras. A modo de ilustración, la Tabla 7 presenta temas de datos cualitativos por dimensiones de la QoS, con indicadores y criterios de evaluación sugeridos.

La Tabla 8 presenta métodos cualitativos para evaluar la QoS y sus fortalezas y limitaciones. Un menú en la Tabla 8 va seguido de una descripción de las consideraciones en torno al análisis.

Tabla 8: Temas e indicadores de datos cualitativos por dimensión de la QoS, con criterios de evaluación

Dimensión	Tema	Indicador	Descripción de la evaluación
Diseño	Tema y plan de investigación	Desafío global/regional	Apropiada, realista
	Diseño	Coherencia, claridad	Apropiada
	Metodología	Integridad, adecuada para el propósito	Rigor, claridad
Insumos	Base de habilidad	Disciplina*	Apropiada
	Composición de equipos	Diversidad, género, disciplina*	Apropiada, inclusiva, multi/transdisciplinariedad
	Infraestructuras	Laboratorios, campos*	Adecuada
	Financiamiento	Compromiso de los donantes*	Adecuada
	Capacitación	Útil para las actividades planificadas	Apropiada, adecuada
Procesos	Alianzas	Inclusión, reconocimiento*	Participación equitativa de los miembros del equipo en el codiseño y la ejecución
	Género	Concienciación, capacidad de respuesta*	Género integrado en el diseño y la implementación
	Funciones y responsabilidades	Definición de funciones y responsabilidades*	Claridad
	Evaluación del desempeño	Incentivos*	Recompensas por la calidad
	Consecuencias negativas	Consecuencias, riesgos	Estrategia de evaluación y mitigación de riesgos
Resultados	Comunicación	Métodos y herramientas*	Relevancia para el público objetivo
	Entorno propicio	Concienciación, comprensión	Posicionamiento y focalización apropiados
	Creación de redes	Compromiso de múltiples partes interesadas*	Adecuada e inclusiva
	Vínculos a políticas	Compromiso de los formuladores de políticas*	Apropiada y dirigida
	Preparación para escalabilidad	Compromiso de múltiples partes interesadas	Contribución a los resultados de desarrollo
	Generación de bienes públicos internacionales (IPG)	Posicionamiento para adopción e impacto*	Amplitud de la aplicabilidad

*Indicadores utilizados en las Revisiones de los CRP 2020.

Tabla 8: Menú de métodos cualitativos sugeridos para la evaluación de la QoS, con fortalezas y limitaciones

Métodos	Definición general	Fortalezas	Limitaciones
Entrevistas/ FGD	Consultas con las principales partes interesadas, ya sea individualmente (entrevistas) o reuniendo a personas de procedencias o experiencias similares para tratar un tema de interés específico en un grupo de discusión focal (FGD).	Útil para recopilar información sobre experiencias, comprender las necesidades satisfechas y no satisfechas y proporcionar ideas para la mejora.	Subjetividad del proceso. Riesgo de omitir a actores importantes debido a asimetrías de poder. Riesgo de sesgos/inexactitudes en el autoinforme y el pensamiento grupal.
Revisión de expertos, revisión documental	Evaluación de la documentación, literatura e informes disponibles.	Revisión documental por parte de SME agrega credibilidad y rigor. Disminución de la presión de tiempo sobre la entidad evaluada tras la entrega de los documentos; inversiones en procedimientos de automatización y minería permiten un análisis de alto rendimiento; es posible que los materiales ya hayan sido revisados por pares o que su calidad haya sido asegurada por terceros, lo que respalda la credibilidad.	Sesgos de expertos. Proceso laborioso (típicamente manual) que requiere muchas horas-persona y experiencia en la materia relevante.
Síntesis de evidencia	Recopilación de información de una variedad de fuentes.	De mayor utilidad para los tomadores de decisiones/formuladores de políticas para la recepción de información resumida; fomenta la observación de tendencias y patrones a lo largo del tiempo y el espacio.	Dependiente de la variedad de evidencia de calidad que se pueda encontrar.
TdC	Descripción e ilustración completas de cómo se espera lograr el cambio deseado en un contexto específico.	Ayuda a identificar si un proyecto cumple con sus objetivos originales; puede ayudar a ajustar los proyectos para satisfacer mejor las necesidades de los beneficiarios finales; podría ser una buena herramienta para la evaluación de la relevancia y el alcance.	Algunas suposiciones pueden ser incorrectas; por tanto, confiar únicamente en la TdC puede conducir a desatender hallazgos desafiantes; es posible que los financiadores no siempre reconozcan su valor.

A continuación, se presentan detalles y consideraciones adicionales sobre los métodos cualitativos seleccionados.

Entrevistas y FGD: entrevistas semiestructuradas son probablemente el método más útil para las revisiones de la QoS. Estas entrevistas se componen de preguntas específicas y enfocadas y también brindan flexibilidad para preguntas abiertas o más exploratorias. Estas entrevistas se pueden hacer a nivel individual con informantes clave o tener lugar en forma de un FGD, enfocado solo en 1-2 temas principales. El FGD es útil cuando hay una pregunta clave que probablemente se responda mejor a través de una discusión estructurada. Personas, a diferencia de documentos, son la «fuente de datos» y deben seleccionarse según criterios de transparencia²¹ p. ej., conocimiento del sector concreto).

Algunos actores clave con los que colaborar podrían ser:

- Líderes en gestión y ciencia del CGIAR (áreas de acción, plataformas de impacto).
- Equipos de intervención: director, jefe de programa, líder de MEA y otros, según corresponda.
- Personal clave, incluidos líderes de proyecto, gerentes.
- Gerentes de investigación de la red de centros de investigación implicados en un proyecto.
- Científicos en los sistemas nacionales de investigación agrícola (NARS).
- Donantes según información adicional del [Cuadro de indicadores de Financiadores](#).
- Socios que incluyen academia, las ONG y el sector privado.
- Investigadores en la fase inicial de su carrera, definidos como investigadores con un doctorado recientemente obtenido.

²¹La investigación cualitativa proporciona protocolos de muestreo específicos que deben cumplirse.

Revisión documental: utilizado para identificar términos de búsqueda clave, patrones y temas que abordan preguntas de evaluación en torno a los resultados, procesos, insumos y diseño/rigor del diseño.

En CGIAR, los documentos clave incluye, pero no se limita a:

- **Informes de iniciativas.** El PRMF 2022-2030 exige que cada iniciativa de CGIAR desarrolle planes de trabajo y presupuesto anuales, monitorice su progreso y publique un informe anual contra los objetivos establecidos y los resultados alcanzados.
- **TdC.** Las iniciativas planifican e informan su progreso anual frente a una TdC que incorpora resultados e indicadores en todas las esferas de control, influencia e interés de la Iniciativa, y se ajusta anualmente mediante un proceso de reflexión.
- **Informes de resultados e impactos** (p. ej., historias de resultados). Informes breves que describen la contribución de los proyectos de investigación a los resultados y el impacto del desarrollo. Un ejemplo de CGIAR son los Informes de casos de resultados e impactos, útiles para comprender la eficacia en términos de resultados y logros a lo largo de las vías de impacto.
- **Estudios de evaluación de impacto.** Las intervenciones y los socios de CGIAR implementan estudios de evaluación de impacto para probar los supuestos en la TdC con el fin de contribuir a su mejora y aumentar su impacto.

5c. Análisis y triangulación de datos

Los métodos seleccionados se acompañarían de diferentes técnicas de análisis y triangulación.

Datos cualitativos: los datos cualitativos extraídos de documentos o entrevistas pueden organizarse de forma manual, p. ej., utilizando Word o Excel, o con software informático. Algunos de los paquetes de software más comunes incluyen Atlas.ti, MAXQDA y NVivo. Si bien la mayoría de los análisis de datos cualitativos son iterativos a lo largo del proceso de recopilación de datos, algunos paquetes de software pueden ayudar con el análisis de los datos extraídos de entrevistas y documentos. Estos incluyen Cynefin SenseMaker, Sprockler y NarraFirma. Es necesaria una formación específica para utilizar estos paquetes de análisis de datos.

Rúbricas: las rúbricas establecen criterios y estándares para diferentes niveles de desempeño y describen y valoran cómo sería el desempeño en cada nivel. En el análisis cualitativo, el uso de rúbricas permite mitigar la subjetividad.²² Las preguntas en la Tabla 4 y el Anexo 7 están formuladas de manera que permite aplicar un sistema sencillo de puntuación, donde rojo indica un problema grave (=No), amarillo un problema menor que se puede resolver (=parcialmente) y verde que la dimensión relacionada con la QoS está funcionando bien para ese indicador/pregunta específico (=Sí).

Tabla 9: Criterios para evaluar la calidad de publicaciones revisadas por pares seleccionadas

Criterios de evaluación	Enfoque de evaluación
¿Representan los resultados (conocimientos presentados en el documento) conocimientos ampliamente aplicables (IPG) relevantes para los objetivos de la intervención?	<p>Escala de puntuación:</p> <p>0=resultados no relevantes para la agricultura y el cambio climático</p> <p>1=sin aplicabilidad más amplia (solo relevancia local)</p> <p>2=aplicabilidad potencialmente más amplia, pero no explicada</p> <p>3=presenta una aplicabilidad más amplia</p> <p>4=aplicabilidad internacional significativa</p>
Calidad (y adecuación) del lugar de publicación	Observación de un lugar de baja calidad o inapropiado en relación con el tema y la calidad del trabajo documental
Coautoría	Observación del alcance de la coautoría, quiénes son y su adecuación
La calidad general de la publicación (incluidos criterios adicionales a discreción del evaluador/SME)	Breve evaluación general (aproximadamente 100-150 palabras)

Métodos mixtos-Evaluación de la calidad de las publicaciones revisadas por pares para relevancia y credibilidad: en las [Revisiones de los CRP 2020](#) (consulte los Anexos para [RTB](#), [CCAFS](#), [WLE](#), y [Livestock](#)), irrevisión exhaustivas de expertos de resultados seleccionados, incluidas publicaciones revisadas por pares y otras publicaciones técnicas y productos físicos (germoplasma, innovaciones digitales y servicios) agregaron credibilidad y rigor para abordar el criterio de evaluación de la QoS y guiar recomendaciones sobre orientaciones futuras. Los criterios combinaron aspectos cuantitativos y cualitativos y utilizaron una escala de puntuación unificada en las doce revisiones. Enfocada específicamente en los resultados, la

Tabla 10 proporciona criterios específicos para evaluar la calidad de las publicaciones revisadas por pares.

Triangulación de datos apoya conclusiones científicas específicas de alta calidad y recomendaciones basadas en la evidencia y mejora la credibilidad de la evaluación. La triangulación de datos puede ocurrir durante todo el proceso de recopilación y análisis de datos, así como al final. La triangulación facilita la validación de datos a través de la verificación cruzada de más de dos fuentes. Pone a prueba la consistencia de los hallazgos obtenidos a través de diferentes instrumentos y aumenta la posibilidad de controlar, o al menos evaluar, algunas de las amenazas o múltiples causas que influyen en nuestros resultados.

²²Un ejemplo de uso de rúbricas de un programa en CIFOR y IUCN <https://www.cifor.org/wp-content/uploads/dfid/KNOWFOR%20-%20Rubrics%20and%20Guidance%20Notes.pdf> del siguiente informe [International Forestry Knowledge Programme \(publishing.service.gov.uk\)](#)

Paso 6: Comunicación de los resultados relacionados con la QoS para mejorar la adopción del aprendizaje y las recomendaciones

La evaluación de la QoS respalda procesos sólidos de toma de decisiones en entornos de investigación para el desarrollo. La comunicación eficaz de los procesos de evaluación de la QoS y los hallazgos y conclusiones relacionados con la QoS es clave para la credibilidad, el aprendizaje y la adopción de las recomendaciones y lecciones de la evaluación. Cuando los usuarios centrales de la evaluación son conscientes del aprendizaje potencial de la evaluación y de sus roles tanto durante el proceso de evaluación como en la adopción posterior de este aprendizaje, es más probable que se actúe sobre las recomendaciones. Para la QoS y, de hecho, todos los tipos de evaluaciones de procesos y desempeño, normalmente hay dos tipos de recomendaciones²³:

- **Recomendaciones formales:** están enumeradas en el apartado «recomendaciones» en un informe de evaluación. Las recomendaciones formales, incluidas las subrecomendaciones, deben recibir una Respuesta de la Gerencia (MR) por escrito, de acuerdo con la Política de evaluación de CGIAR (2022). Una MR bien documentada es un documento de aprendizaje que puede contribuir a ayudar a CGIAR a evitar y mitigar problemas estratégicos, políticos o sistémicos que surjan en la programación futura.
- **Recomendaciones informales:** en lugar de una recomendación formal, un equipo de evaluación puede decidir hacer una «sugerencia» u observar una «lección aprendida». Hay muchas razones para esto, incluso cuando la recomendación excede el alcance de sus ToR, es necesario priorizar recomendaciones más sustantivas o es probable que la recomendación no sea factible o procesable. Para tales recomendaciones informales, si bien no es necesario, la gerencia puede optar por dar una respuesta.

En particular, cuando se aplica el criterio de la QoS, es importante equilibrar las recomendaciones independientes por dimensiones de la QoS e incorporar las demás recomendaciones que pueden haber surgido de la evaluación sobre eficiencia, sostenibilidad y otros criterios de evaluación.

Independientemente del tipo de recomendaciones, es muy recomendable que la evidencia sintetizada por las dimensiones relacionadas con la QoS se presente a los gerentes científicos, investigadores y otras partes interesadas relevantes durante la fase de validación y se detalle en el informe de evaluación final. Posteriormente, la comunicación efectiva aumentará la comprensión de las evaluaciones de procesos y desempeño y generará confianza en las partes interesadas para motivar respuestas positivas a las recomendaciones y lecciones aprendidas.

Una evaluación puede resultar en múltiples productos de comunicación dirigidos a diferentes grupos de usuarios. La identificación de las necesidades específicas de información de los grupos clave determinará el tipo apropiado de productos de comunicación y formas de gestionar el conocimiento evaluativo (KM). Esto puede incluir presentaciones, vídeos cortos o podcasts, blogs, resúmenes, presentaciones en conferencias y otros medios para llegar a las audiencias con contenido oportuno y apropiado, para facilitar la entrega de mensajes y el uso de la información.²⁴

²³Consultar las Guías de evaluación del CGIAR: Compromiso y respuesta de la gerencia.

²⁴Consultar las Guías de evaluación del CGIAR: Informe de evaluación final.

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, 9(1), 215824401982957. <https://doi.org/10.1177/2158244019829575>

Bryman, A. (2004). Triangulation and measurement. Retrieved from Department of Social Sciences, Loughborough University, Loughborough, Leicestershire

Carneiro, B., Resce, G., Läderach, P., Schapendonk, F., & Pacillo, G. (2022). What is the importance of climate research? An innovative web-based approach to assess the influence and reach of climate research programs. *Environmental Science & Policy*, 133, 115-126. <https://doi.org/10.1016/j.envsci.2022.03.018>

Carter, N. (2014, September). The use of triangulation in qualitative research. In *Oncology nursing forum* (Vol. 41, No. 5, p. 545). 10.1188/14.ONF.545-547

CAS Secretariat (CGIAR Advisory Services Shared Secretariat). (2020). CGIAR Research Program 2020 Reviews: Policies, Institutions, and Markets. Rome. CAS Secretariat Evaluation Function. <https://cas.cgiar.org/> and [List of Annexes](#)

CAS Secretariat (CGIAR Advisory Services Shared Secretariat). (2020). CGIAR Research Program 2020 Reviews: Roots, Tubers and Bananas (RTB). Rome. CAS Secretariat Evaluation Function. <https://cas.cgiar.org/>

CAS Secretariat (CGIAR Advisory Services Shared Secretariat). (2021). Synthesis of Learning from a Decade of CGIAR Research Programs. Rome. CAS Secretariat Evaluation Function. <https://cas.cgiar.org/>

CGIAR Independent Advisory and Evaluation Service (2022). CGIAR Evaluation Framework. Rome: CGIAR Independent Advisory and Evaluation Service.

CGIAR System Management Office. (2020). CGIAR Performance and Results Management Framework 2022-2030. Montpellier. CGIAR. <https://hdl.handle.net/10568/113793>

CGIAR System Organization. (2021). CGIAR 2030 Research and Innovation Strategy: Transforming food, land, and water systems in a climate crisis. Montpellier. CGIAR System Organization. <https://cgspace.cgiar.org/handle/10568/110918>

CGIAR System Organization (2021). 2022-24 Investment Prospectus: pooling funds for research and innovation to transform food, land, and water systems. Montpellier. CGIAR System Organization. <https://www.cgiar.org/research/investment-prospectus/>, https://storage.googleapis.com/cgiarorg/2021/06/Document-SC13_02_Endorsed-2022-24-Investment-Prospectus.pdf and [Companion Document](#).

De Col, V., Jani, S., Rünzel, M., Tobon, H., Almanzar, M., See, D. S., & Bonaiuti, E. (2021). Case Study on the Monitoring-Quality Assurance Processor-API: A Tool to Support CGIAR Quality Assurance Process for Peer-reviewed Publications. Beirut, Lebanon: International Center for Agricultural Research in the Dry Areas (ICARDA). <https://repo.mel.cgiar.org/handle/20.500.11766/66480>

De Col, V. (2022). Alone we can do so little; together we can do so much. CAS Secretariat. <https://cas.cgiar.org/evaluation/news/alone-we-can-do-so-little-together-we-can-do-so-much>

Hcéres (2022). Evaluated entities area. <https://www.hceres.fr/en/evaluated-entities-area>

Heale, R., & Forbes, D. (2013). Understanding triangulation in research. *Evidence-based nursing*, 16(4), 98-98. <http://dx.doi.org/10.1136/eb-2013-101494>

IDRC (International Development Research Centre). (2022). The International Development Research Centre's Research Quality Plus (RQ+) Assessment Instrument. www.idrc.ca/RQplus

IEA (Independent Evaluation Arrangement) (2015). CGIAR Standards for independent external evaluation. <https://iaes.cgiar.org/sites/default/files/pdf/Standards.pdf>

IEA (Independent Evaluation Arrangement). (2015) IEA Workshop on Evaluating Quality of Science. Report. Rome 10-11 December 2015 https://iaes.cgiar.org/sites/default/files/pdf/Report_QoSWorkshop-final-1.pdf

Jappe, A. (2020). Professional standards in bibliometric research evaluation? A meta-evaluation of European assessment practice 2005–2019. PLOS One, 15(4), e0231735. <https://doi.org/10.1371/JOURNAL.PONE.0231735>

Langfeldt, L., & Scordato, L. (2015). Assessing the Broader Impacts of Research: A Review of Methods and Practices. Oslo: Nordic Institute for Studies in Innovation, Research and Education. <https://nifu.brage.unit.no/nifu-xmlui/handle/11250/282742>

Neylon, C., Willmers, M., & King, T. (2014). Rethinking impact: Applying Altmetrics to southern African research. Paper/Scholarly Communication in Africa Programme. Canada. IDRC. <http://hdl.handle.net/10625/53461>

OECD/DAC. (2019). Better criteria for better evaluation: Revised evaluation criteria definitions and principles for use. The Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC). <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>

Science-Metrix & CGIAR Advisory Services Secretariat Evaluation Function (2022). Bibliometric Analysis to Evaluate Quality of Science in the Context of One CGIAR. Technical Note. Rome. <https://cas.cgiar.org/evaluation/publications/bibliometric-analysis-evaluate-quality-science-context-one-cgiar>

Science-Metrix. (2018). Review of the Human Frontier Science Program Final Report. Quebec. Science-Metrix inc. <https://www.hfsp.org/node/12547#book/>

Serrat, O. (2017). Social network analysis. In Knowledge solutions (pp. 39-43). Springer, Singapore. [Social Network Analysis | SpringerLink](#)

Wilson, V. (2014). Research methods: Triangulation. Evidence-based library and information practice, 9(1), 74-75. <https://doi.org/10.18438/B8WW3X>

Anexo 1: partes interesadas clave para el desarrollo de directrices

El desarrollo de las directrices siguió un enfoque consultivo, inclusivo e iterativo que incluyó debates estratégicos y operativos con partes interesadas tanto internas como externas del CGIAR, incluidos, entre otros temas, la gobernanza y gestión de CGIAR, la MELCOP de CGIAR y la revisión por pares externa. El Anexo 1 proporciona una lista de los expertos que participan en las actividades realizadas. La base fundamental y las fuentes de conocimiento

provinieron de un [taller sobre la evaluación de la calidad de la ciencia](#) (2015), [EvalForward²⁵ discussion](#) (EN, FR, ES) sobre la evaluación de la ciencia, tecnología e innovación en un contexto de desarrollo (2022) y la participación de expertos en la [Conferencia de la Sociedad Europea de Evaluación](#) (junio 2022); y de recomendaciones seleccionadas de la bibliometría sobre la utilización de métodos mixtos en evaluaciones de la [Nota técnica](#).²⁶

Grupo de partes interesadas	Taller de validación, 23 junio 2022	EvalForward ²⁷ debate de la CoP (EN, FR, ES)		
Financiadores	<ul style="list-style-type: none"> - Raphael Nawrotzki, responsable de M&E del Fondo internacional para la investigación agrícola (FIA) de la sociedad alemana para la Cooperación Internacional (GIZ), Alemania 			
CGIAR Independent Science for Development Council (ISDC)	<ul style="list-style-type: none"> - Holger Meinke, presidente de ISDC; también profesor investigador adjunto de sostenibilidad alimentaria global en la Universidad de Tasmania, Australia - Andrew Ash, miembro de ISDC. Director y jefe de AJ Ash and Associates. Profesor adjunto en la Escuela de agricultura y ciencias de la alimentación, Universidad de Queensland, Australia - Amy Beaudreault, directora de la Secretaría de ISDC, IAES, Italia - Pierre Boulanger, asesor de la Secretaría de ISDC, IAES, Italia 			
Socios de investigación entre otros, universidades	<ul style="list-style-type: none"> - Claudio Proietti, asesor de M&E en el CIRAD, Francia - Nobert Tchouaffe, investigador del Instituto panafricano para el desarrollo, Camerún - Rachid Serraj, director asociado de estrategia de la Universidad Politécnica Mohammed VI, Marruecos - Richard Tinsley, profesor emérito de la Universidad Estatal de Colorado, EE. UU. - Valeria Pesce, facilitadora de alianzas en el GFAR, Italia 			
<ul style="list-style-type: none"> - Etienne Vignola Gagné, analista de Science-Metrix/Elsevier, coautor de la Nota técnica, Canadá 				
Grupo de referencia de evaluación (ERG) del Independent Advisory and Evaluation Service de CGIAR	<ul style="list-style-type: none"> - Guy Poppy, director y profesor en la Universidad de Southampton, Reino Unido <table border="1" data-bbox="403 1496 1495 1693"> <tr> <td data-bbox="403 1496 874 1693"> Zenda Ofir, científica y evaluadora internacional a tiempo completo - comentarios escritos, Suiza/Sudáfrica </td> <td data-bbox="874 1496 1495 1693"> <ul style="list-style-type: none"> - Sonal D Zaveri, fundadora y coordinadora en GENSA, comunidad de evaluadores de Asia Meridional, India - Ola Ogunyinka, especialista en monitoreo, evaluación e impacto del NRI, Universidad de Greenwich, Reino Unido </td> </tr> </table>		Zenda Ofir, científica y evaluadora internacional a tiempo completo - comentarios escritos, Suiza/Sudáfrica	<ul style="list-style-type: none"> - Sonal D Zaveri, fundadora y coordinadora en GENSA, comunidad de evaluadores de Asia Meridional, India - Ola Ogunyinka, especialista en monitoreo, evaluación e impacto del NRI, Universidad de Greenwich, Reino Unido
Zenda Ofir, científica y evaluadora internacional a tiempo completo - comentarios escritos, Suiza/Sudáfrica	<ul style="list-style-type: none"> - Sonal D Zaveri, fundadora y coordinadora en GENSA, comunidad de evaluadores de Asia Meridional, India - Ola Ogunyinka, especialista en monitoreo, evaluación e impacto del NRI, Universidad de Greenwich, Reino Unido 			
CGIAR	<ul style="list-style-type: none"> - Alessandra Furtado, directora interina de la PCU; jefa de gestión de proyectos en el CIP, Mozambique - Valentina De Col, responsable de sistemas de información agrícola en el ICARDA, Alemania - Graham Thiele, exdirector del CRP en RTB, Perú 			
<ul style="list-style-type: none"> - Bia Carneiro, especialista en investigación social y medios de CGIAR FOCUS seguridad climática, Portugal 				

²⁵EvalForward es una CoP en torno a la seguridad alimentaria, la agricultura y el desarrollo rural. Reúne a funcionarios y profesionales para intercambiar experiencias y fortalecer las capacidades de evaluación a nivel de país. EvalForward pretende contribuir a la evaluación del progreso hacia el ODS 2: poner fin al hambre, lograr la seguridad alimentaria y la mejora de la nutrición y promover la agricultura sostenible.

²⁶Science-Metrix se integró con el equipo de Research Analytics and Data Services (RADS) de Elsevier en 2018.

²⁷EvalForward es una CoP en torno a la seguridad alimentaria, la agricultura y el desarrollo rural.

Anexo 1: partes interesadas clave para el desarrollo de directrices cont.

Grupo de partes interesadas	Taller de validación, 23 junio 2022	EvalForward ²⁷ debate de la CoP (EN, FR, ES)
Independiente y privado	<ul style="list-style-type: none"> - Beverly Parsons, directora ejecutiva de InSites; presidente de la Asociación estadounidense de evaluación (2013-2015), EE. UU. - Sara Vaca, consultora de visualización de datos, CGIAR y UNICEF, Francia 	<ul style="list-style-type: none"> - Keith Child, consultor de evaluación de IAES, anteriormente experto en MEA para el CRP de CGIAR sobre agua, tierra y ecosistemas (WLE), Canadá - Lennart Raetzell, gerente de Syspons GmbH, Alemania
Expertos independientes, Revisiones evaluativas del CRP de 2020	<ul style="list-style-type: none"> - Jillian Lenne - Consultora independiente y editora. Previamente SME para la QoS en las Revisiones de los CRP 2020 (RTB y GLDC), Reino Unido - Donna Podems - Evaluadora para 2020 CRP WHEAT (trigo), Sudáfrica - John Morton - Profesor de antropología de desarrollo en el NRI; SME para la revisión del CRP CCAFS 2020, Reino Unido 	<ul style="list-style-type: none"> - Paolo Sarfatti, asesor senior de evaluación estratégica y técnica, Italia - Paul Engel, SME para la revisión del CRP PIM 2020; líder de equipo en conocimiento, perspectivas e innovaciones, Países Bajos
La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO)	<ul style="list-style-type: none"> - Rachel Sauvinet Bedouin, oficial jefe de evaluación; anteriormente jefa del Independent Evaluation Arrangement (IEA) de CGIAR - comentarios por escrito 	<ul style="list-style-type: none"> - Nanae Yabuki, oficial de evaluación - Ibtissem Jouini, especialista en evaluación regional para el Oriente Próximo y norte de África (RNE); anteriormente consultora de evaluación en IAES - Serdar Bayryyev, oficial jefe de evaluación
CGIAR Independent Advisory and Evaluation Service (IAES)	<ul style="list-style-type: none"> - Svetlana Negroustoueva, responsable de la Función de evaluación - Allison Grove-Smith, directora del IAES - Gaia Gullotta, consultora analista de datos - Inese Berzina, coordinadora administrativa 	

Entrevistas a informantes clave	Tema	
CGIAR	<ul style="list-style-type: none"> - Bia Carneiro, especialista en investigación social y medios, anteriormente en el CRP en CCAFS - Valentina De Col, responsable de sistemas de información agrícola, ICARDA 	Análisis de redes sociales (SNA)
Universidad de Curtin, Australia	<ul style="list-style-type: none"> - Cameron Neymar, profesor de comunicación de investigación, Universidad de Curtin, Centro de Cultura y Tecnología 	Altmetrics
Universidad de Ciudad del Cabo, Sudáfrica	<ul style="list-style-type: none"> - Michelle Willmers, Gerente de publicación e implementación del proyecto Libros de texto digitales en abierto para el desarrollo 	
Fondo de población de las Naciones Unidas (UNPFA)	<ul style="list-style-type: none"> - Lamin Massaquoi, experto en datos y líder del equipo de África Oriental - Kais Al-Abhar, analista de monitoreo y evaluación 	Uso del marco RQ+, Centro internacional de investigaciones para el desarrollo (IDRC)

²⁷EvalForward es una CoP en torno a la seguridad alimentaria, la agricultura y el desarrollo rural.

Annex 2: criterios de evaluación del CGIAR²⁸

Además del criterio de evaluación de la QoS, hay disponibles otras guías más detalladas sobre otros criterios en la Red de evaluación del desarrollo (EvalNet) de la OECD/DAC <https://www.oecd.org/desarrollo/evaluacion/>.

Calidad de la ciencia (QoS): el criterio evaluativo de la QoS se refiere a la credibilidad y legitimidad científicas. La definición del criterio deriva del marco de referencia de la QoR4D, que registra a nivel de todo el CGIAR System Organization el acuerdo sobre la naturaleza y evaluación de la calidad de la investigación. La QoR4D describe la calidad de la investigación de acuerdo con cuatro elementos clave: relevancia, credibilidad científica, legitimidad y eficacia.²⁹ La relevancia y la eficacia son tratados por separado como criterios de evaluación.

Relevancia: la medida en que los objetivos y el diseño de la intervención responden a las necesidades, políticas y prioridades de los usuarios/clientes, socios e instituciones a nivel mundial, regional y nacional y continúan haciéndolo incluso si las circunstancias cambian. De acuerdo con el marco de la QoR4D, se presta atención a la importancia, el significado y la utilidad del trabajo implementado en el contexto del problema, asociado con la capacidad del CGIAR para abordar los problemas.

Eficacia: la medida en que la intervención logra, y/o se espera que logre, sus objetivos y resultados, incluidos los resultados diferenciales entre subgrupos de usuarios/clientes. En coherencia con el marco de la QoR4D y en el contexto del CGIAR, este criterio considera la medida en que la investigación está posicionada para su uso y ha generado conocimiento, productos y servicios con alto potencial para abordar un problema y contribuir a las innovaciones, los resultados y los impactos. Por tanto, la eficacia implica que la investigación ha sido diseñada, implementada y posicionada para su uso dentro de una teoría dinámica del cambio, con liderazgo apropiado, capacitación, diversidad de habilidades de investigación y apoyo al entorno propicio para traducir el conocimiento en prácticas y ayudar a generar los resultados deseados.

Coherencia: la compatibilidad de la intervención con otras intervenciones en un país, sector o dentro de CGIAR; su ajuste general. La coherencia interna aborda las sinergias y los vínculos entre la intervención y otras intervenciones realizadas dentro del CGIAR, y la coherencia de la intervención con las normas y los estándares internacionales pertinentes a los que se adhiere CGIAR. La coherencia externa considera la consistencia de la intervención con las intervenciones de otros actores en el mismo contexto, es decir, su complementariedad, armonización y coordinación con otros, su valor añadido y su evitación de la duplicación de esfuerzos.

Eficiencia: la medida en que la intervención aporta o es probable que aporte resultados de manera económica y oportuna; es decir, el uso general de los recursos. «Económica» se refiere a la conversión de insumos (fondos, experiencia, recursos naturales, tiempo, etc.) en productos, resultados e impactos de la manera más rentable posible, en comparación con alternativas factibles en el contexto. La entrega «oportuna» significa que se hace dentro del plazo previsto, o un plazo razonablemente ajustado a las demandas del contexto en evolución. Este criterio puede incluir la evaluación de la eficiencia operativa (en qué medida se gestionó bien la intervención).

Sostenibilidad: la medida en que los beneficios netos de la intervención tengan continuidad o es probable que continúen. Este criterio se centra en la continuación de los beneficios, no en la financiación externa, y destaca la naturaleza multidimensional de la sostenibilidad.

Impacto: la medida en que la intervención ha generado o se espera que contribuya a generar efectos significativos positivos o negativos, intencionados o no intencionados, de mayor nivel. El impacto aborda el resultado último y los efectos potencialmente transformadores de la intervención.

²⁸ <https://iaes.cgiar.org/evaluation/publications/cgiar-evaluation-policy>

²⁹ Unas pautas codiseñadas sobre la evaluación de la QoS en el CGIAR detallan el enfoque y los métodos para poner en práctica el criterio de evaluación de la QoS de esta Política.

Annex 3: ejemplos de preguntas y subpreguntas de evaluación por métodos y fuentes de datos

Pregunta de evaluación	Ejemplos de subpreguntas de evaluación	Métodos y fuentes de datos
Diseño		
EQ1. ¿El diseño de la investigación es adecuado para los desafíos de desarrollo en el contexto?	<ol style="list-style-type: none"> 1. ¿Existe un vínculo documentado entre un objetivo declarado y las áreas de impacto? 2. ¿Existe evidencia a nivel de programa u otra evidencia de cambio de metodología si los objetivos de la investigación cambiaron? Aprendizaje: ¿se describen e integran claramente los resultados/hallazgos de investigaciones anteriores? 3. ¿Son las preguntas y la metodología de investigación adecuadas para el propósito y están alineadas con el problema de investigación? 4. ¿Cómo de innovador es la investigación y la ciencia? ¿Se tuvo en cuenta la ventaja comparativa? 5. ¿Cómo de interconectados están el diseño de la investigación y los ODS, incluso dentro de cada ODS? 6. ¿Cómo de alineado está el diseño de la investigación con las prioridades compartidas, de múltiples financiadores y de los socios? 7. ¿Cómo y qué socios participaron en el codiseño de la ciencia entregada? 8. ¿Está el vínculo entre el plan MEA/IA y los indicadores en la TdC claramente definido en el diseño de la iniciativa de investigación? 	Propuestas de iniciativas e informes; informes de ISDC de revisión de la iniciativa; entrevistas; TdC (original y revisiones); revisiones de propuestas ex ante de ISDC
Insumos		
EQ2a. ¿Hasta qué punto eran los insumos necesarios adecuados y suficientes para entregar los resultados y productos planificados? (se relaciona con EQ.2 en la Figura 4)	<ol style="list-style-type: none"> 1. ¿La composición de los equipos de ejecución de la investigación era suficientemente diversa (inclusiva en términos de género, edad/jóvenes investigadores y nacionalidad)? 2. ¿Hubo un rango apropiado de disciplinas y habilidades, dado el tema de la investigación? 3. ¿Las infraestructuras físicas de investigación (p. ej., laboratorios, parcelas experimentales, etc.) eran adecuadas? 4. ¿El financiamiento de la investigación era suficiente y oportuna? 5. ¿La capacitación del equipo de investigación y los socios abordó las necesidades con respecto al trabajo planificado, incluidos los aspectos no científicos? 	Informes de iniciativas; bibliometría; entrevistas; SNA; visitas de campo/laboratorio de expertos; revisiones presupuestarias versus los planes; evaluaciones de necesidades y registros de capacitación
Procesos		
EQ2b. ¿En qué medida el proceso de gestión aseguró la QoS, incluida la credibilidad y la legitimidad científicas de la investigación y las operaciones? (se relaciona con EQ.2 en la Figura 4)	<ol style="list-style-type: none"> 1. ¿Cuál fue el nivel de confianza, entendimiento y compromiso con los socios (de diferentes tipos)?³⁰ 2. ¿En qué medida se integraron los socios externos en el equipo de investigación y sus operaciones? 3. ¿Había políticas establecidas para la ética de la investigación; se implementaron bien? 4. ¿Los roles y las responsabilidades se definieron claramente y se implementaron según lo planeado, en línea con las esferas de control e influencia de las TdC? 5. ¿Había políticas establecidas para mecanismos internos de revisión por pares, para mejorar el aprendizaje? 6. ¿Había políticas establecidas para el mentoring y la formación del personal investigador junior? 7. ¿Se establecieron estrategias de evaluación de riesgos y mitigación? 	Entrevistas; FGD; informes de iniciativas; actas de reuniones; análisis interno de políticas; análisis de la matriz de riesgos

³⁰Alineación con un «Marco de alianzas»

Anexo 3: ejemplos de preguntas y subpreguntas de evaluación por métodos y fuentes de datos cont.

Pregunta de evaluación	Ejemplos de subpreguntas de evaluación	Métodos y fuentes de datos
Resultados		
EQ3. ¿Cómo contribuyen los resultados de la intervención al avance de la ciencia?	1. ¿Calidad y cantidad de publicaciones científicas y técnicas y otros productos?	Bibliometría; Altmetrics; informes de iniciativas; entrevistas; TdC; FGD; análisis de expertos de publicaciones científicas (calidad de la revisión documental)
¿De qué manera los resultados de la investigación son de alta calidad?:	2. ¿Cuántas publicaciones se produjeron? ¿Cuál fue el factor de impacto de las revistas? ¿Cuál fue la proporción de publicaciones muy citadas? ¿Quiénes eran los autores más productivos?	
	3. ¿Se citaron las publicaciones a través de diferentes canales (en lugar de los más tradicionales) como blogs, Twitter, etc.?	
a. Productos físicos: germoplasma, innovaciones y servicios digitales	4. ¿Se comunicaron/difundieron claramente los hallazgos de la investigación y los resultados relacionados?	
	5. ¿Hubo una solicitud de los socios y/u otras partes interesadas para presentar la investigación y sus resultados? ¿Cuál fue el alcance de las publicaciones en los países focales y en los NARS?	
	6. ¿Cómo contribuyen los resultados a las innovaciones basadas en la ciencia, las capacitaciones específicas y el asesoramiento sobre políticas?	
	7. ¿Se han documentado las vías y las contribuciones de los resultados a los ODS?	
	8. ¿Cómo reflejan los resultados las preocupaciones en torno a GDI, el medio ambiente o los esfuerzos de localización/adaptación a contextos específicos, por ejemplo? Coautores del sur global (prevalencia).	
	9. ¿Pueden estos productos tener aplicabilidad más amplia y potencial de impacto a escala?	
	b. Publicaciones de investigación y técnicas, materiales de capacitación, guías, mecanismos de apoyo para la toma de decisiones y asesoramiento sobre políticas	
		Informes de iniciativas; entrevistas; FGD; análisis de expertos

Anexo 4: parámetros y análisis de datos

Revisiones de documentos: Bel Cuadro de texto 1 a continuación muestra documentación clave para la revisión documental en la evaluación de la QoS además de la utilización de otros criterios de evaluación (ver la guía [Elaboración y uso de los análisis de evaluabilidad](#)).

Cuadro de texto1: ejemplos de documentación y fuentes externas

1. Documentos de propuestas o estrategia
2. Teoría del cambio
3. Marco de resultados u otros documentos con un desglose de aportaciones, actividades y productos, resultados e impactos esperados
4. Listados de proyectos (con su documentación relacionada)
5. Listados de personas de contacto para partes interesadas internas/externas e informantes clave
6. Evaluaciones, estudios y evaluaciones de impacto previos o independientes
7. Evaluaciones de impacto
8. Bases de datos clave con información potencialmente relevante
9. Publicaciones revisadas por pares
10. Resúmenes de políticas
11. Documentos de trabajo
12. Preimpresiones
13. Informes técnicos

Altmetrics rastrea diversas fuentes para cotejar conversaciones sobre investigaciones que ocurren en línea todos los días. Aporta métricas y datos cualitativos complementarios a las métricas tradicionales basadas en citas. Pueden incluir (sin limitarse a) revisiones de pares en Faculty of 1000, citas en Wikipedia y en documentos de política pública, discusiones en blogs de investigación, cobertura de los principales medios, marcadores de gestores bibliográficos como Mendeley y menciones en redes sociales como Twitter. Utilizando fuentes de Internet, Altmetrics puede revelar mucho sobre la frecuencia con la que se discuten y utilizan a nivel mundial los artículos de revistas y otros productos académicos, como conjuntos de datos. Es útil hacer un seguimiento e informar sobre la atención recibida por una obra a través de canales diferentes a los más comunes. Altmetrics tiene el potencial de alcanzar una reconceptualización integral de lo que califica como impacto, lo que debería ser recompensado en las estructuras institucionales de incentivos y recompensas, y cómo rastrear y promover la participación con las alianzas de la sociedad civil (Neylon, 2014). La revisión evaluativa del programa de investigación CGIAR [Políticas, instituciones y mercados \(PIM\)](#) (2020) describe el uso de la puntuación de atención de Altmetrics³¹ en la evaluación de la QoS (consulte el [Anexo](#)). La puntuación de atención de Altmetrics se obtiene de forma gratuita por el equipo de PRMS y no requiere que las iniciativas o los centros paguen una suscripción anual, a menos que deseen utilizar los servicios avanzados proporcionados por Altmetrics para su propio uso.

Análisis de redes sociales (SNA) es una forma gráfica de representar el número y la robustez de las conexiones entre personas, incluidos investigadores, instituciones, socios gubernamentales, etc. El SNA busca comprender las redes y sus participantes y tiene dos enfoques principales: los actores y sus interrelaciones en un contexto social específico (Serrat 2017). Un ejemplo como estudio de caso es el uso del SNA en el programa de investigación sobre cambio climático, agricultura y seguridad alimentaria (CCAFS) del CGIAR (Carneiro [2022]). Los autores readaptan de manera efectiva los datos disponibles públicamente de fuentes digitales, como redes sociales y sitios web, empleando minería de textos y SNA para evaluar la influencia y el alcance del programa entre las partes interesadas a diversos niveles. Es más, el siguiente blog en los CRP de RTB y GLDC describe el potencial del SNA para complementar otros métodos y métricas utilizados para M&E de la QoS. El ejemplo en coautoría de los centros del CGIAR ICARDA y CIP con la Universidad de Florida muestra cómo integrar diferentes dimensiones como el enfoque geográfico, el género y la duración de las colaboraciones y la bibliometría en el SNA aporta valor al análisis de los insumos y procesos de la QoS, bajo la QoR4D. Programas informáticos recomendados para SNA incluyen: VOSviewer, un software para construir y visualizar redes bibliométricas (de código abierto y gratuito); Gephi, un software de visualización y exploración para diferentes tipos de gráficos y redes (de código abierto y gratuito). Este tipo de análisis requiere la cooperación entre un experto en el campo y un evaluador interno que pueda interpretar los resultados teniendo en cuenta el contexto de CGIAR.

³¹It is a score for a specific research output that indicates the amount of attention it has received. The score is derived from an automated algorithm and represents a weighted count of the amount of attention received for research output.

Annex 4: parámetros y análisis de datos cont.

Bibliometría incluye un poderoso conjunto de herramientas para evaluar el desempeño científico de diversas entidades (países, regiones, instituciones o investigadores) mediante la caracterización de las dimensiones de sus productos científicos (es decir, en su mayoría publicaciones científicas revisadas por pares), como el tamaño de su producción, sus patrones de colaboración, su impacto científico y la medida en que recombinan diferentes ámbitos de conocimiento a través, entre otras cosas, alianzas con expertos de un conjunto diversificado de ámbitos. La bibliometría se utiliza para evaluar el financiamiento, las políticas y las actividades científicas -en particular para evaluar los resultados de esas intervenciones sobre la excelencia en la investigación- y una amplia gama de organismos gubernamentales/no gubernamentales la implementan y utilizan para este fin a nivel internacional (Jappe, 2020). Los indicadores bibliométricos también se pueden utilizar para evaluar otras dimensiones, como los procesos e insumos; consulte el Anexo 6 para conocer las preguntas de evaluación relacionadas con los insumos y procesos que pueden abordarse, al menos parcialmente, mediante bibliometría. Todos los indicadores bibliométricos tienen algunas debilidades cuando se consideran de forma individual. Para la evaluación cuantitativa, un uso extendido de la bibliometría es esencial y puede ser alimentado y mejorado por insumos cualitativos. En consecuencia, es importante utilizar varias líneas de evidencia para triangular los resultados.

La lista de indicadores de alta prioridad informados en la Tabla 6 incluye indicadores de: participación igualitaria de género; proporción de publicaciones que son copublicaciones académico-privadas; integración interdisciplinaria de las ciencias sociales y humanidades (SSH) dentro de las publicaciones; impacto normalizado de citación; interdisciplinaria; y copublicación Sur-Sur, Sur-Norte, alineación temática con los ODS. Excepto por algunos indicadores del impacto normalizado de citación e indicadores interdisciplinarios, las fórmulas de los indicadores son una división simple, expresada como una proporción o porcentaje de un conjunto total de publicaciones. Los indicadores pueden calcularse reuniendo un conjunto general de publicaciones (p. ej., todas las publicaciones de una Iniciativa 2022–24 o todas las publicaciones de un Área de acción). El número de publicaciones en este conjunto es el denominador en las fórmulas del indicador. El numerador se determina contando el número de publicaciones dentro del conjunto general que cumplen con un criterio; p. ej., publicaciones que tengan al menos una mujer coautora o un coautor del Sur o que han recibido al menos una mención periodística según los registros de las bases de datos de Altmetrics. La inclusión en el recuento del numerador también se puede basar en múltiples

criterios; p. ej., la publicación tiene una autora como primera, última o autora de correspondencia y las autoras constituyen el 50% o más de las autorías dentro de la publicación. La priorización de indicadores no indica el nivel de autoridad de un solo indicador frente a otros indicadores. Siempre es recomendable utilizar un panel de indicadores complementarios para capturar diferentes aspectos incluso de un solo fenómeno. El Anexo 7 de la [Nota técnica](#) incluye un listado más amplio de indicadores a tener en cuenta para la evaluación, en función de su viabilidad y de las necesidades de la evaluación.

Web of Science: plataforma de acceso de pago que brinda acceso a múltiples bases de datos que contienen referencias y citas de revistas académicas, actas de congresos y otros documentos de diversas disciplinas académicas. Útil para calcular indicadores bibliométricos y Altmetrics. Para el análisis bibliométrico, se recomienda utilizar [Bibliometrix](#), una herramienta R para el análisis del mapeo de la ciencia y [Biblioshiny](#), una aplicación Shiny para Bibliometrix (ambas gratuitas).

Datos de monitoreo del desempeño: [cuadro de indicadores de resultados de CGIAR](#): El cuadro de indicadores brinda acceso a datos en masa y a publicaciones e Informes de casos de impacto de los resultados (OICR) con información detallada. Creada en sus orígenes con fines de gestión y rendición de cuentas, se espera que esta herramienta evolucione como herramienta de M&E. La información que contiene ha superado un control de calidad. El uso del cuadro de indicadores a través de MEA y MARLO combinados en CLARISA varía de un CRP a otro. [CLARISA](#): CLARISA (Arquitectura del sistema interoperable de resultados agrícolas a nivel CGIAR) es un servicio web que ayuda a transformar los datos brutos sobre la investigación y las actividades del CGIAR en información significativa que puede dar forma a cómo trabajamos y revelar cuáles son nuestros impactos en el desarrollo sobre la reducción de la pobreza, la mejora de la seguridad alimentaria y nutricional para la salud y la mejora de los recursos naturales y los servicios ecosistémicos. En términos de programación, CLARISA es una REST-API, lo que significa que es un tipo de servicio web que permite que los sistemas informáticos trabajen juntos a través de Internet. CLARISA permite que sistemas como MARLO y MEA, entre otros, se comuniquen entre sí, descubre puntos en común entre sus datos y genera información agregada y estandarizada en el idioma requerido para los informes a nivel del Sistema. Funciona mediante el uso de listados de control de términos clave estandarizados, como los utilizados comúnmente por el Marco de estrategia y resultados del CGIAR y los ODS.

Annex 5: mapeo de los criterios de revisión de propuestas y elementos QoR4D frente a los estándares y criterios del marco de evaluación

Estándares del marco de evaluación (CGIAR, 2022) <i>Evaluación ex post del desempeño/procesos</i>	Calidad de la investigación para el desarrollo QoR4D en la práctica para One CGIAR (2021) <i>Criterios de revisión ex ante de propuestas</i>	Marco de referencia de la QoR4D (2020) <i>Elementos de la QoR4D</i>	Criterios de la política de evaluación (CGIAR, 2022) <i>Evaluación ex post de desempeño/procesos</i>
<p>1. <i>Relevancia, uso, y utilidad:</i> todas las evaluaciones son aplicables a la(s) pregunta(s) en cuestión y están diseñadas de manera receptiva y oportuna para su uso en la toma de decisiones, la rendición de cuentas y los procesos de aprendizaje.</p>	<p>1. Problema de investigación claramente definido que aborda las áreas de impacto, es una alta prioridad en las geografías objetivo, está bien alineado con las prioridades compartidas de múltiples financiadores, y está bien informado por los hallazgos de investigaciones anteriores.</p>	<p>Relevancia Eficacia</p>	<p>Relevancia Coherencia Eficacia</p>
<p>2. <i>Independencia, ausencia de sesgo:</i> las evaluaciones inspiran confianza entre todos los usuarios de que la evaluación es lo más objetiva posible con los más altos estándares éticos y códigos de conducta; imparcial, con un sistema establecido contra conflictos de intereses; y ausente de sesgos desde el punto de vista operativo y analítico.</p>	<p>8. Se tiene en cuenta la ética, incluidas las alianzas equitativas, la divulgación de información, los sesgos y los posibles conflictos de intereses; la propuesta define cómo se buscarán/otorgarán las aprobaciones formales de ética de la investigación.</p>	<p>Legitimidad Credibilidad</p>	<p>QoS</p>
<p>3. <i>Transparencia:</i> los procesos (incluidos los métodos) y resultados se divulgan de forma transparente, son rastreables y accesibles por el público.</p>	<p>14. El gasto es justificado y está vinculado de forma transparente a los resultados esperados de la investigación para el desarrollo.</p> <p>15. Se describen los resultados esperados de la investigación (conocimiento, avances técnicos o institucionales, tecnologías o productos específicos, análisis de políticas) y los conocimientos y las brechas a ser abordados son evidentes con un enfoque demostrado en la calidad, la visión de futuro, la relevancia del impacto y su disseminación. Los protocolos para el cumplimiento de datos abiertos y acceso abierto son evidentes en el plan (incluido el presupuesto).</p>	<p>Legitimidad Eficacia</p> <p>Credibilidad Eficacia</p>	<p>QoS Eficacia Eficiencia</p>
<p>4. <i>Legitimidad y participación:</i> las evaluaciones incluyen informantes relevantes y utilizan procesos consultivos para preparar los términos de referencia y la matriz de evaluación. Estos procesos aseguran la calidad del proceso, e incluso, cuando sea apropiado y factible, son representativos de los usuarios finales e intermedios de los resultados de la evaluación.</p>	<p>2. Evidencia mediante la firma conjunta de que la Iniciativa está impulsada por la demanda de las partes interesadas y los socios principales (grupos asesores de inversión, gobiernos, sector privado, financiadores) y colaboradores de investigación, tanto dentro como fuera del CGIAR.</p> <p>11. Las declaraciones de capacidad explican la idoneidad de los proponentes como implementadores del trabajo. Se manifiesta la propuesta de valor y justifica la capacidad e idoneidad del CGIAR para liderar el trabajo. Esto incluye las habilidades, la diversidad y la multi/transdisciplinariedad del equipo de investigación y los colaboradores.</p>	<p>Relevancia Eficacia</p>	<p>Relevancia Coherencia Eficacia QoS Sostenibilidad</p>
<p>5. <i>Capacidad de respuesta</i> a GDI: el diseño y la ejecución de la evaluación, la creación de equipos, y la presentación de informes se esfuerzan por abordar los parámetros de GDI en su totalidad. Las evaluaciones tendrán en cuenta quienes son los involucrados y los beneficiarios del trabajo.</p>	<p>9. El diseño de la investigación y la implementación propuesta demuestran la inclusión social y de género que serán rastreables en los resultados.</p>	<p>Legitimidad Eficacia</p>	<p>Relevancia Eficacia QoS Impacto</p>
<p>6. <i>Ética y equidad:</i> las evaluaciones tienen en cuenta cuestiones de ética en la investigación y los resultados e integran consideraciones éticas y de equidad en el diseño y la implementación de la evaluación.</p>	<p>8. Se tiene en cuenta la ética, incluidas las alianzas equitativas, la divulgación de información, los sesgos y los posibles conflictos de intereses; la propuesta define cómo se buscarán/otorgarán las aprobaciones formales de ética de la investigación.</p>	<p>Legitimidad Credibilidad</p>	<p>Relevancia Eficacia QoS</p>

Anexo 5: mapeo de los criterios de revisión de propuestas y elementos QoR4D frente a los estándares y criterios del marco de evaluación cont.

Estándares del marco de evaluación (CGIAR, 2022) <i>Evaluación ex post del desempeño/procesos</i>	Calidad de la investigación para el desarrollo QoR4D en la práctica para One CGIAR (2021) <i>Criterios de revisión ex ante de propuestas</i>	Criterios de la política de evaluación (CGIAR, 2022) <i>Evaluación ex post de desempeño/procesos</i>
<p>7. <i>Evaluabilidad</i>: la evaluabilidad hace referencia a la medida en que una intervención puede evaluarse de forma fiable y creíble; el concepto es crucial a una cultura basada en los resultados. Un enfoque sólido en la evaluabilidad durante la fase de diseño facilita, en general, la medibilidad, el monitoreo y la evaluación posterior.</p>	<p>15. Se describen los resultados esperados de la investigación (conocimiento, avances técnicos o institucionales, tecnologías o productos específicos, análisis de políticas) y son evidentes los conocimientos y las brechas que serán abordados a través de un enfoque demostrado en la calidad, la visión de futuro, la relevancia del impacto y la diseminación. Los protocolos para el cumplimiento de datos abiertos y acceso abierto son evidentes en el plan (incluido el presupuesto).</p> <p>16. El plan M&E para la Iniciativa está claramente definido, con flexibilidad de adaptación. El plan de M&E apoya una gestión y aprendizaje efectivos, incluida la recopilación de datos de línea de base y los procesos de evaluación y revisión correspondientes para la toma de decisiones en cada etapa y para la implementación de correcciones en la ejecución actual incluso la recopilación de datos a nivel basal, y los procesos de evaluación y revisión correspondientes para la toma de decisiones en cada etapa y para la implementación de correcciones en la ejecución actual. El M&E ocurre durante la vida de la Iniciativa y se aplica de manera proactiva para reflexionar sobre la Teoría de Cambio TdC y, cuando corresponda, adaptarla.</p>	<p>Credibilidad Eficacia</p> <p>Eficacia</p> <p>Eficiencia</p> <p>QoS</p> <p>Credibilidad Eficacia</p> <p>Legitimidad</p>
<p>8. <i>Credibilidad y robustez</i>: los métodos empleados son creíbles y replicables. La calidad de una evaluación depende de la competencia profesional y metodológica de los evaluadores y de la utilización de datos fiables y triangulados.</p>	<p>5. La metodología y los métodos de investigación (y las actividades de apoyo) son adecuados para su propósito, factibles, vanguardistas y rigurosos en la recopilación y el análisis de datos, y las limitaciones se han establecido con claridad.</p>	<p>Credibilidad Relevancia</p> <p>Eficacia</p> <p>QoS</p>
<p>9. <i>Medibilidad</i>: métodos sólidos sustentan la medibilidad y la replicabilidad. En la medida de lo posible, las evaluaciones miden el desempeño del CGIAR mediante métodos cuantitativos y/o cualitativos. La medibilidad proporciona comparabilidad entre marcos temporales, grupos o teorías alternativas.</p>	<p>3. Las preguntas de investigación, los objetivos, los productos y los resultados están alineados con el problema de investigación, son medibles con hitos y etapas bien definidos y susceptibles de evaluación y acción correctiva a lo largo del ciclo de vida del proyecto.</p>	<p>Relevancia Eficacia</p> <p>Coherencia</p> <p>Eficacia</p>
<p>10. <i>Rendición de cuentas mutua</i>: en CGIAR, las expectativas de evaluación se combinan con inversiones adecuadas en recursos financieros y humanos necesarios. La capacidad y los sistemas para la recopilación de datos y la información en tiempo real respaldan la rendición de cuentas mutua.</p>	<p>13. Los mecanismos de gestión de proyectos y (en su caso) las medidas de gobernanza y supervisión científica adicionales respaldan de manera eficaz y eficiente los objetivos de la Iniciativa.</p> <p>14. El gasto es justificado y está vinculado de forma transparente a los resultados esperados de la investigación para el desarrollo.</p>	<p>Legitimidad Credibilidad</p> <p>QoS</p> <p>Eficiencia</p> <p>Legitimidad Eficiencia</p>
<p>11. <i>Eficiencia</i>: la evaluación evita duplicaciones, costos o redundancias innecesarias en relación a otras valoraciones evaluativas.</p>	<p>6. Análisis de compensaciones y sinergias en las áreas de impacto de CGIAR; la evaluación ex ante de los beneficios del proyecto proporciona una justificación lógica para escalar los impactos.</p> <p>14. El costeo es justificado y está vinculado de forma transparente a los resultados esperados de la investigación para el desarrollo.</p>	<p>Eficacia Credibilidad</p> <p>Coherencia</p> <p>Eficacia</p> <p>Eficiencia</p> <p>Impacto</p> <p>Legitimidad Eficiencia</p>

Anexo 6: matriz de diseño de evaluación, criterios de la QoS

Dimensión	Subpregunta de evaluación	Elementos a evaluar	Criterios de evaluación	Tipo de método	Fuentes de datos, métodos, análisis	Política de evaluación	
QoS_EQ1: ¿El diseño de la investigación es adecuado para los desafíos de desarrollo en el contexto?							
Diseño	1.1. ¿La investigación que subyace al objetivo de la intervención se alinea con las prioridades compartidas de múltiples financiadores?	Relevancia de la investigación	Relevancia y coherencia de la agenda de investigación	Cualitativo	Informes de intervenciones de la teoría de cambio TdC; entrevistas; rúbricas	Relevancia	
	1.2. ¿Se ha evaluado y documentado sistemáticamente la ventaja comparativa?	Relevancia y coherencia de la investigación		Mixto	Análisis de compensaciones	Relevancia Coherencia	
	1.3. ¿El vínculo entre el plan de evaluación de impacto y los indicadores en la TdC está claramente definido en el diseño de la investigación?	Diseño de la investigación	Rigor de la investigación experimental		Cualitativo	Informes de intervenciones de la TdC; entrevistas; rúbricas	Relevancia QoS
	1.4. ¿Cómo de interconectados están el diseño de la investigación y los ODS, incluso dentro de cada ODS?	Relevancia de la investigación	Alineación con los ODS		Cualitativo/ cuantitativo	Informes de intervenciones de la TdC; bibliometría; rúbricas	Relevancia
	1.5. ¿El diseño de la investigación de la intervención es apropiado y está articulado con claridad?	Diseño de la investigación	Rigor de la investigación experimental		Cualitativo	Informes de intervenciones; entrevistas	Relevancia QoS
	1.6. ¿Son adecuados la metodología y los métodos de investigación para el propósito de la intervención?	Diseño de la investigación	Rigor de la investigación experimental		Cualitativo	Informes de intervenciones; entrevistas	Relevancia QoS
	1.7. ¿Un objetivo definido de una intervención aborda las áreas de impacto de CGIAR?	Relevancia de la investigación	Relevancia y coherencia de la agenda de investigación		Cualitativo	Informes de intervenciones de la TdC; entrevistas	Relevancia
QoS_EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima?							
Insumos	2.1. ¿Hasta qué punto los insumos necesarios eran adecuados y suficientes para entregar los productos y resultados planificados?	Composición de los equipos de investigación	Adecuación de habilidades y disciplinas científicas; multi/transdisciplinariedad del nivel de integración; inclusión en relación con la diversidad de edad, género y nacionalidad	Cuantitativo Cualitativo	Informes de intervenciones; bibliometría; entrevistas; SNA	QoS	
		Atractivo de los equipos de investigación	Atractivo de los miembros del equipo	Cualitativo	Entrevistas; perfil de los miembros del equipo; informes de intervenciones	QoS	
	2.2. ¿Se describen claramente los resultados/hallazgos de investigaciones anteriores y se integran los aprendizajes?	Reputación de la unidad de investigación	Reputación científica de la unidad; reconocimiento obtenido a través del éxito en convocatorias competitivas para proyectos	Cualitativo	Entrevistas; informes de intervenciones	QoS	
		Financiación	Adecuación y previsibilidad; compromiso de los donantes	Cuantitativo Cualitativo	Entrevistas; informes de intervenciones	QoS	

Anexo 6: matriz de diseño de evaluación, criterios de la QoS cont

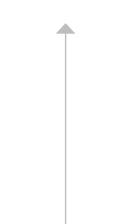
Dimensión	Subpregunta de evaluación	Elementos a evaluar	Criterios de evaluación	Tipo de método	Fuentes de datos, métodos, análisis	Política de evaluación
QoS_EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima? <small>cont.</small>						
Inputs	2.1. ¿Hasta qué punto los insumos necesarios eran adecuados y suficientes para entregar los productos y resultados planificados? 2.2. ¿Se describen claramente los resultados/hallazgos de investigaciones anteriores y se integran los aprendizajes?	Infraestructuras de investigación	Adecuación de laboratorios y campos	Cualitativo	Informes de intervenciones; entrevistas	Eficiencia QoS
		Capacitación	Apropiado y adecuado, útil para las actividades planificadas	Cualitativo	Entrevistas; informes	Eficiencia QoS
		Ventaja comparativa	Mejor conocimiento disponible, identificando los costos relativos de los entregables clave entre las organizaciones identificadas, incluido el CGIAR	Cualitativo	Informes	Coherencia Eficiencia Eficacia
		Relevancia de la investigación	El diseño de la investigación es apropiado y se basa en abordar las brechas evidentes; revisión adecuada y completa de la literatura	Cualitativo	Informes de intervenciones de la TdC; informes; entrevistas	Relevancia Eficacia
		Ventaja comparativa	Mejor conocimiento disponible, identificando los costos relativos de los entregables clave entre las organizaciones identificadas, incluido el CGIAR	Cualitativo	Informes	Coherencia Eficiencia Eficacia
		Participación de las partes interesadas	Partes interesadas apropiadas involucradas en la etapa correcta	Cualitativo	Informes de intervenciones; entrevistas; grupos de discusión focal FGD; análisis de redes sociales SNA	QoS Coherencia
Processes	2.2. ¿En qué medida el proceso de gestión aseguró la QoS, incluida la credibilidad científica y la legitimidad de la investigación y las operaciones?	Alianzas	Enfoque de múltiples partes interesadas; confianza, comprensión y compromiso mutuos; reconocimiento claro de las perspectivas, necesidades, roles y contribuciones de los socios, ventaja comparativa, incluida la resiliencia ³²	Cualitativo	Informes de intervenciones; entrevistas; FGD; SNA; rúbricas	QoS Coherencia
		Ética de la investigación	Políticas establecidas para la ética de la investigación y su implementación	Cualitativo	Informes de intervenciones; entrevistas; grupos de discusión focal FGD; análisis de redes sociales SNA	QoS

³²Independent Science for Development Council. 2022. Identifying and Using CGIAR's Comparative Advantage. [Identificación y uso de la ventaja comparativa del CGIAR]. Rome: CGIAR Independent Advisory and Evaluation Service. <https://iaes.cgiar.org/sites/default/files/pdf/ISDC-Technical-Note-Identifying-and-Using-CGIAR-Comparative-Advantage.pdf>

Anexo 6: matriz de diseño de evaluación, criterios de la QoS cont

Dimensión	Subpregunta de evaluación	Elementos a evaluar	Criterios de evaluación	Tipo de método	Fuentes de datos, métodos, análisis	Política de evaluación		
QoS_EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima? <small>cont.</small>								
Procesos	2.2. ¿En qué medida el proceso de gestión aseguró la QoS, incluida la credibilidad científica y la legitimidad de la investigación y las operaciones?	Compromiso con el conocimiento local	Las comunidades, partes interesadas o poblaciones locales participaron efectivamente y se tuvieron en cuenta en el proceso de investigación	Cualitativo	Informes de intervenciones; entrevistas; grupos de discusión focal FGD	QoS		
		Funciones y responsabilidades	Funciones y responsabilidades claramente definidas	Cualitativo Cuantitativo	Revisión documental; encuesta; entrevistas; FGD	QoS		
		Mecanismos de revisión interna	Políticas establecidas para los mecanismos de revisión interna y su implementación	Cualitativo	Informes de intervenciones; entrevistas; grupos de discusión focal FGD	QoS		
		Tutoría y formación de personal junior	Políticas establecidas para tutoría y formación del personal junior y para su implementación	Cualitativo	Informes de intervenciones; entrevistas; FGD	QoS		
		Género	GDI en la implementación	Cualitativo Cuantitativo	Entrevistas; FGD; cuadro de indicadores de GDI	QoS		
		Evaluación del desempeño	El trabajo de calidad es recompensado	Cualitativo	Entrevistas; FGD; encuestas	Efectividad QoS		
		Gestión de riesgos	Establecimiento de estrategias de evaluación de riesgos y mitigación	Cualitativo	Informes de auditoría interna; entrevistas	Eficiencia QoS		
		Protocolos para el cumplimiento de datos abiertos y acceso abierto (incluido el presupuesto)	Accesibilidad de datos e información	Cualitativo Cuantitativo	Bibliometría; entrevistas			
		QoS_EQ3: ¿Cómo contribuyen los resultados de la intervención al avance de la ciencia?						
		Resultados	3.6. ¿De qué manera los resultados de la investigación, como variedades mejoradas, herramientas de conocimiento y publicaciones, son de alta calidad?	Calidad y cantidad de publicaciones científicas y técnicas	Número de publicaciones; índice H de autores más productivos; factor de impacto de las revistas; proporción de publicaciones altamente citadas (HCP); índice de distribución de citas (CDI); promedio de citación relativa (ARC); integración multidisciplinar; puntuaciones de Altmetrics	Cuantitativo	Altmetrics; bibliometría; informes de intervenciones	QoS

Anexo 6: matriz de diseño de evaluación, criterios de la QoS cont

Dimensión	Subpregunta de evaluación	Elementos a evaluar	Criterios de evaluación	Tipo de método	Fuentes de datos, métodos, análisis	Política de evaluación
QoS_EQ3: ¿Cómo contribuyen los resultados de la intervención al avance de la ciencia? <small>cont.</small>						
Resultados  	3.7. ¿Los resultados se reflejaron en nuevas políticas y/o contribuyeron a la sociedad donde se busca el cambio?	Vínculos a políticas	Políticas que citan los productos de investigación; trabajo en red a nivel gubernamental; estudios de impacto elaborados; participación de los formuladores de políticas	Cuantitativo	Informes (incluidas las evaluaciones de impacto, en caso de disponibilidad); entrevistas; SNA; TdC	QoS Relevancia Eficacia
	3.8. ¿Los productos físicos, p. ej., variedades mejoradas e innovación digital, eran de alta calidad y relevantes para los usuarios de la próxima etapa?	Desarrollo de productos físicos, p. ej., variedades mejoradas e innovaciones digitales	Aplicabilidad más amplia; adaptabilidad del producto físico al contexto; preparación para escalabilidad; relevancia para el público objetivo	Cualitativo Cuantitativo	Informes de intervenciones; entrevistas; FGD; TdC	Eficacia Sostenibilidad
	3.9. ¿Se comunicaron claramente los resultados de la investigación?	Comunicación de los resultados de la investigación	Relevancia para los públicos objetivos	Cualitativo	Informes de intervenciones; entrevistas	Relevancia

Anexo 7: uso de rúbricas para la evaluación cualitativa de la QoS

Dimensión	Pregunta de evaluación	Elementos a evaluar	Criterios de evaluación	No	Parcialmente	Sí
QoS_EQ1: ¿El diseño de la investigación es adecuado para los desafíos de desarrollo en el contexto?						
Design	1.1. ¿La investigación que subyace al objetivo de la intervención se alinea con las prioridades compartidas de múltiples financiadores?	Relevancia de la investigación	Relevancia y coherencia de la agenda de investigación	La investigación no se alinea con las prioridades compartidas de múltiples financiadores	La investigación se alinea solo parcialmente con las prioridades compartidas de múltiples financiadores	La investigación se alinea por completo con las prioridades compartidas de múltiples financiadores
	1.2. ¿Se ha evaluado y documentado sistemáticamente la ventaja comparativa?	Relevancia y coherencia de la investigación		La ventaja comparativa no fue evaluada y documentada	Existe evidencia parcial, no sistemática, sobre la ventaja comparativa	La intervención tiene una evaluación sistemática claramente documentada de la ventaja comparativa existente de la intervención en el contexto
	1.3. ¿El vínculo entre el plan de evaluación de impacto y los indicadores en la TdC está claramente definido en el diseño de la investigación?	Diseño de la investigación	Rigor de la investigación experimental	El diseño de la investigación no define claramente la relación entre el estudio y la TdC	Se menciona la TdC, pero no se define claramente su vínculo con el plan de evaluación de impacto en el diseño de la investigación	El vínculo entre el plan de evaluación de impacto y los indicadores en la TdC está claramente definido en el diseño de la investigación
	1.4. ¿Cómo de interconectados están el diseño de la investigación y los ODS, incluso dentro de cada ODS?	Relevancia de la investigación	Alineación con los ODS	El diseño de la investigación no se alinea con los ODS	El diseño de la investigación se alinea con al menos un ODS	El diseño de la investigación claramente se alinea con un mayor número de los ODS y también es relevante indirectamente para otros ODS
	1.5. ¿El diseño de la investigación de la intervención es apropiado y está articulado con claridad?	Diseño de la investigación	Rigor de la investigación experimental	No hubo un diseño de investigación claramente articulado	El diseño de la investigación era articulado, pero dejó algunos vacíos	El diseño de la investigación era apropiado y claramente articulado
	1.6. ¿La metodología y los métodos de investigación son adecuados para el propósito de la intervención?	Diseño de la investigación	Rigor de la investigación experimental	El diseño de la investigación no cumplió con los estándares metodológicos y no era adecuado para el propósito	Se logró parcialmente el cumplimiento de los estándares metodológicos	La metodología y los métodos de investigación son adecuados para el propósito
	1.7. ¿Un objetivo definido de una intervención aborda las áreas de impacto de CGIAR?	Diseño de la investigación	Rigor de la investigación experimental	Las preguntas de investigación no están claramente formuladas y/o no están alineadas con el problema de investigación	Las preguntas de investigación están solo parcialmente alineadas con el problema de investigación	Las preguntas de investigación están claramente formuladas y abordan adecuadamente el problema de investigación
	3.4. ¿Un problema de investigación definido aborda las áreas de impacto?	Relevancia de la investigación	Relevancia y coherencia de la agenda de investigación	El problema de investigación no se relaciona directamente con ningún área de impacto	El problema de investigación se refiere a áreas de impacto, pero el vínculo no está claramente definido	El problema de investigación definido aborda claramente las áreas de impacto

Anexo 7: uso de rúbricas para la evaluación cualitativa de la QoS cont.

Dimensión	Pregunta de evaluación	Elementos a evaluar	Criterios de evaluación	No	Parcialmente	Sí
QoS_EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima?						
Insumos	2.1. ¿Hasta qué punto eran adecuados y suficientes los insumos necesarios para entregar los productos y resultados planificados?	Composición de los equipos de investigación	Adecuación de habilidades y disciplinas científicas; nivel de integración multi/transdisciplinariedad	Bajo nivel de integración. Falta de diversidad en términos de habilidades y baja integración disciplinar	Nivel adecuado de integración en términos de habilidades y disciplinas científicas, pero hay espacio para mejora	Alto nivel de integración en términos de habilidades y disciplinas científicas
		Composición de los equipos de investigación	Inclusividad en relación con diversidad de edad, género y nacionalidad	Bajo nivel de integración. Falta de diversidad dentro de los equipos en relación a género, edad y/o nacionalidad	Nivel adecuado de integración, pero hay espacio para mejoras adicionales	Alto nivel de integración y diversidad dentro de los equipos de investigación
		Atractivo de los equipos de investigación	Atractivo de los miembros del equipo	Los miembros del equipo no son atractivos por su calibre	Los miembros del equipo tienen buenos perfiles, pero les falta algo de experiencia para ser atractivos del todo	Alto calibre de los miembros de los equipos, alto atractivo
		Atractivo de la unidad de investigación	Reputación científica de la unidad; reconocimiento obtenido a través del éxito en convocatorias competitivas para proyectos	La unidad no tiene un historial de ganar convocatorias competitivas para proyectos	La unidad tiene un historial de ganar algunas convocatorias competitivas de proyectos (entre 1 y 5)	La unidad es conocida por haber ganado varias convocatorias competitivas de proyectos (más de 5)
		Financiación	Adecuación y previsibilidad; compromiso de los donantes	Financiamiento insuficiente. Escaso compromiso de los donantes	Buen nivel, pero no predecible, de financiación. El compromiso de los donantes es parcial	El financiamiento es adecuado y predecible. Los donantes tienen un alto nivel de compromiso
		Infraestructuras de investigación	Adecuación de laboratorios y campos	Laboratorios y campos son inadecuados para el propósito de la investigación	Laboratorios y campos son parcialmente adecuados. Algunas mejoras son necesarias	Laboratorios y campos son adecuados para realizar las investigaciones
		Capacitación	Apropiada y adecuada, útil para las actividades planificadas	Hubo una falta de capacitación a lo largo del proyecto	Hubo algunas actividades de capacitación, pero con un enfoque limitado	Hubo numerosas y buenas actividades de capacitación a lo largo del proyecto
	2.2. ¿Se describen e integran claramente los resultados/hallazgos de la investigación?	Relevancia de la investigación	El diseño de la investigación es apropiado y se basa en abordar las brechas evidentes; revisión adecuada e íntegra de la literatura	La revisión de la literatura/documental, cuando evidente, era insuficiente y en gran medida desactualizada	La revisión de la literatura/documental era apropiada pero no completamente exhaustiva o no completamente integrada	La revisión de la literatura/documental era apropiada y exhaustiva

Anexo 7: uso de rúbricas para la evaluación cualitativa de la QoS cont.

Dimensión	Pregunta de evaluación	Elementos a evaluar	Criterios de evaluación	No	Parcialmente	Sí
QoS_EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima? cont'd						
Procesos	2.2.¿En qué medida el proceso de gestión aseguró la QoS, incluyendo la credibilidad científica y la legitimidad de la investigación y las operaciones?	Alianzas	Confianza mutua, comprensión y compromiso; reconocimiento claro de las perspectivas, las necesidades, los roles y las contribuciones de los socios; enfoque de las múltiples partes interesadas	La relación con los socios no estaba clara. Era difícil encontrar alineaciones de necesidades con los socios. Débil contribución de los socios	Buen compromiso de algunos socios y con diferencias en el compromiso a lo largo de la duración de la intervención	La relación con los socios ha sido buena durante toda la duración del proyecto. Las perspectivas, las necesidades, los roles y las contribuciones de los socios siempre eran claras
		Ética de la investigación	Políticas establecidas para la ética en la investigación y su implementación	No hay políticas establecidas para la ética en la investigación y/o las políticas establecidas no se respetaron	Las políticas establecidas para la investigación no eran exhaustivas y/o se aplicaron solo parcialmente	Las políticas establecidas para la investigación eran exhaustivas y se aplicaron de manera amplia
		Compromiso con el conocimiento local	Las comunidades locales, partes interesadas o poblaciones tuvieron un compromiso efectivo y se tuvieron en cuenta durante el proceso de investigación	El compromiso con los contextos apropiados se ha descuidado durante el proceso de investigación	Los contextos y el compromiso se han tenido en cuenta durante el proceso de investigación	El compromiso con las comunidades locales, partes interesadas o poblaciones ocurrió de forma adecuada y creíble
		Funciones y responsabilidades	Funciones y responsabilidades claramente definidas	Funciones y responsabilidades no estaban claramente definidas	Funciones y responsabilidades parcialmente claras. Algunos aspectos se consideraron poco claros durante la investigación	Funciones y responsabilidades estaban claramente definidas a lo largo de la duración de la investigación
		Mecanismos de revisión interna	Políticas establecidas para los mecanismos de revisión interna y su implementación	Ausencia de políticas para los mecanismos de revisión internos de revisión	Políticas existen, pero no son exhaustivas/no se aplican en su totalidad	Políticas establecidas y aplicadas plenamente durante todo el proceso de la investigación
		Tutoría y formación de personal junior	Políticas establecidas para tutoría y formación de personal junior y para su implementación	Inexistencia de políticas para tutoría y formación de personal junior	Políticas establecidas para tutoría y formación de personal junior, pero necesitan mejorarse	Buenas políticas establecidas para tutoría y formación de personal junior. Amplia implementación de las políticas
		Género	Existencia de GDI en la implementación	No se tuvo en cuenta el equilibrio de género y roles en el equipo de investigación	La consideración de género era limitada en la composición y las funciones del equipo de investigación	Se hizo hincapié en el equilibrio de género y en la adecuación de los roles en el equipo de investigación

Anexo 7: uso de rúbricas para la evaluación cualitativa de la QoS cont.

Dimensión	Pregunta de evaluación	Elementos a evaluar	Criterios de evaluación	No	Parcialmente	Sí
QoS_EQ2: ¿Los insumos y procesos son adecuados para generar ciencia creíble y legítima? <small>cont.</small>						
Procesos	2.2. To what extent did the management process ensure the Quality of Science, including scientific credibility, and legitimacy, of the research and operations?	Evaluación del desempeño	El trabajo de calidad es recompensado	No existe un sistema de evaluación y recompensa del desempeño	Existencia de un sistema de evaluación del desempeño, pero el trabajo de calidad no siempre ha sido recompensado	Realización de evaluación del desempeño durante el proceso de investigación y recompensas por el trabajo de calidad
		Gestión de riesgos	Establecimiento de estrategias de evaluación y mitigación de riesgos	Ausencia de estrategias de evaluación y mitigación de riesgos	Identificación de estrategias de evaluación y mitigación de riesgos, pero no siempre se implementaron	Buen uso de la evaluación de riesgos e implementación de estrategias de mitigación en caso necesario
		Protocolos para el cumplimiento de datos abiertos y acceso abierto (incluido el presupuesto)	Accesibilidad de datos e información	Falta de datos e información	Disponibilidad solo parcial de datos e información	Disponibilidad completa de datos e información. Presencia de protocolos para el cumplimiento
EQ3: ¿Cómo contribuyen los resultados de la intervención al avance de la ciencia?						
Resultados	3.1. ¿Los productos físicos, p. ej., variedades mejoradas e innovación digital, eran de alta calidad y relevantes para los usuarios de la próxima etapa?	Desarrollo de productos físicos, p. ej., variedades mejoradas e innovaciones digitales	Aplicabilidad más amplia; adaptabilidad del producto físico al contexto; preparación para escalabilidad; relevancia para el público objetivo	Los productos no se aplican ampliamente y no están preparados para escalabilidad	Los productos se pueden aplicar parcialmente. La escalabilidad de los productos no es fácil	Aplicabilidad del producto. El producto está preparado para escalabilidad. Es relevante para el público objetivo
	3.2. ¿Se comunicaron claramente los resultados de la investigación?	Comunicación de los resultados de la investigación	Relevancia para los públicos objetivos	No se comunican bien los resultados de la investigación	Los resultados de la investigación se comunican, pero no se utilizan todos los medios para llegar a las audiencias objetivo	Los resultados de la investigación se comunican bien y los públicos objetivo son fácilmente alcanzados

Anexo 8: bibliometría - glosario e indicadores prioritarios

ID: tanto una referencia alfabética a la dimensión de relevancia de la QoR4D como un número único.

Título del indicador: nombre del indicador.

Implementación: modalidad de implementación (por quién y cuándo):

- CGIAR +: podría implementarse a nivel interno por la PPU, la Comunidad de práctica MEA o analistas comprometidos con el CAS por recomendación de Science-Metrix en el futuro.
- Externo: tendría que implementarse por un proveedor externo en el futuro.
- Piloto: indicador aún en fase de diseño; puede implementarse por la PPU, la Comunidad de práctica MEA o analistas comprometidos con el CAS o proveedores externos, pero en todos los casos requiere algo de I+D, sin garantía de éxito.

Tiempo: número de años tras la conclusión de un proyecto durante los cuales pueden evaluarse las publicaciones generadas a través de ese proyecto (considerando que las publicaciones relevantes aún se publican en los dos años inmediatamente posteriores al último año formal de un proyecto).

Limitaciones: una tipología de limitaciones genéricas. Incluye lo siguiente:

- No normalizado: el indicador no está normalizado en la actualidad o nunca podrá normalizarse para controlar los sesgos de campo y las tendencias anuales.
- Limpieza: requiere esfuerzos sustanciales para armonizar los metadatos.
- Óptimo desconocido: el conocimiento actual no permite determinar por completo una buena práctica en la dimensión medida por este indicador; puntuaciones altas en la medición pueden tener efectos adversos en las prácticas de investigación.
- Proxy imperfecto: el indicador captura solo un componente reducido de un fenómeno de interés más amplio.
- Posible captura de tokenismo: los indicadores cuantitativos de equidad entre grupos típicamente no capturan una equidad plenamente alcanzada, sino solo las manifestaciones externas de equidad. Esta limitación se superpone con la limitación de proxy imperfecto.
- Definición categórica compleja: la asignación de un producto a una categoría puede basarse en juicios o pautas necesariamente imperfectas.
- Errores de metadatos: hay deficiencias reconocidas en los metadatos que suelen utilizarse para calcular este indicador, ya sea porque los propios autores de la publicación cometen errores o porque la codificación y el análisis de las bases de datos bibliográficas son imperfectos.
- Discrepancias entre planes y logros: puede haber grandes diferencias entre las propuestas de proyectos y la ejecución de proyectos.

Anexo 8: bibliometría - glosario e indicadores prioritarios cont.

Tabla 10: SEjemplo de matriz de colecta de datos (DCM) para indicadores bibliométricos de alta prioridad³³

ID	Título	Implementación	Tiempo (años)	Limitaciones
L23	Proporción de publicaciones con participación de mujeres en la autoría	CGIAR+	+3	No captura el equilibrio o la equidad; posible captura de tokenismo; software de pago (Namsor); margen de error (especialmente para nombres asiáticos)
L24	Proporción de publicaciones que logran el equilibrio de género en la autoría clave	CGIAR+	+3	Software de pago (NamSor); margen de error (especialmente para nombres asiáticos)
L26	Proporción de copublicaciones Norte-Sur/Sur-Sur	CGIAR+	+3	No normalizado; limpieza; óptimo desconocido; proxy imperfecto; no captura el equilibrio o la equidad
L27	Participación de autores del Sur como primer o último autor o autor de correspondencia	CGIAR+	+3	Tasa de error en los datos de afiliación; proxy imperfecto (equidad Sur-Norte)
L31	Visualización de diagrama de cuerdas de copublicaciones internacionales	Piloto	+3	Errores de metadatos (datos de afiliación); base de conocimientos limitada (indicador novedoso); proxy imperfecto (equidad en la integración multinacional)
R31	Alineación temática con el tema relevante para los ODS	CGIAR+	+3	Proxy imperfecto (transferencia de conocimiento para el desarrollo); base de conocimientos limitada; errores de metadatos
R34	Proporción de copublicaciones académicas-privadas	CGIAR+ o externa	+3	Normalización difícil; limpieza extensiva; definición categórica compleja; proxy imperfecto (transferencia de tecnología)
R38	Proporción de publicaciones altamente citadas (HCP)	Externa	+5	Proxy imperfecto (calidad de publicación y logro intelectual); obligación de 30 o más publicaciones; computable 2 años o más después del año de publicación
R39	Índice de distribución de citas (CDI)	Externa	+5	Proxy imperfecto (calidad de publicación y logro intelectual); obligación de 30 o más publicaciones; computable 2 años o más después del año de publicación
R41	Media de citas relativas (ARC)	CGIAR+	+5	Proxy imperfecto (calidad de publicación y logro intelectual); sensible a valores atípicos; obligación de 30 o más publicaciones; computable 2 años o más después del año de publicación
R42	Índice de integración interdisciplinar	Externa	+3	Proxy imperfecto (integración disciplinaria intelectual); sesgo hacia la interdisciplinariedad novedosa y radical; índice abstracto es lo más significativo como parte de las comparaciones
R43	Proporción de publicaciones con alta interdisciplinariedad	Externa	+3	Proxy imperfecto (integración disciplinaria intelectual); sesgo hacia la interdisciplinariedad novedosa y radical

³³Bibliometric Analysis to Evaluate Quality of Science in the Context of One CGIAR | CAS | CGIAR Advisory Services

Anexo 8: bibliometría - glosario e indicadores prioritarios cont.

Tabla 10: Ejemplo de matriz de colecta de datos (DCM) para indicadores bibliométricos de alta prioridad (cont.)

ID	Título	Implementación	Tiempo (años)	Limitaciones
R44	Índice de integración multidisciplinar	Externa	+3	Proxy imperfecto (integración disciplinaria colaborativa); sesgo hacia la diversidad disciplinaria novedosa y radical
R45	Proporción de publicaciones con alta multidisciplinariedad	Externa	+3	Proxy imperfecto (integración disciplinaria colaborativa); sesgo hacia la diversidad disciplinaria novedosa y radical
R46	Visualización del diagrama de cuerdas de la interdisciplinariedad (en particular para capturar la integración entre ciencias sociales y humanidades)	Externa	+3	Proxy imperfecto (integración interdisciplinaria); sesgo hacia la interdisciplinariedad novedosa y radical



Independent
Advisory and
Evaluation
Service

Contacto

CGIAR Independent Advisory and Evaluation Service (IAES)

Via di San Domenico 1, 00153 Rome, Italy

Email: IAES@cgiar.org

URL: <https://iaes.cgiar.org/evaluation>

Sigue la «Independent Advisory and Evaluation Service»
en las redes sociales

