

Guía: Datos abiertos sobre *CAMBIO CLIMÁTICO*



Guillermo Durán
FELLOW ESTADO ABIERTO 2018



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional.

Creative Commons Atribución/Reconocimiento 4.0 Licencia Pública Internacional — CC BY 4.0

Título: Guía: Datos abiertos sobre cambio climático

Autor: Guillermo Durán

Edición: Javiera Atenas, Daniel Villatoro, Juan Manuel Casanueva y Fabrizio Scrollini.

Enero 2019. San José, Costa Rica

DOI: <http://doi.org/10.5281/zenodo.2560546>

Diseño: Daniel Villatoro

Citar como: Durán, G. (2019). Guía: Datos Abiertos sobre Cambio Climático. (J. Atenas, D. Villatoro, J. M. Casanueva, & F. Scrollini, Eds.). San José de Costa Rica: Escuela de Datos e Iniciativa Latinoamericana por los Datos Abiertos. <http://doi.org/http://doi.org/10.5281/zenodo.2560546>

contenidos

02	¿Qué es una guía de datos abiertos?
03	¿Qué son los datos abiertos?
04	¿Qué características deben tener los datos abiertos?
06	Componente legal de los datos abiertos en Costa Rica
09	¿Cuál es la importancia de los datos abiertos?
11	¿Qué son los estándares de datos abiertos?
12	Tipos de estándares de datos
14	¿Qué son las licencias abiertas y por qué son importantes?
15	¿Qué son los formatos abiertos de datos?
18	¿Cómo se evalúan los portales de datos abiertos?
20	Evaluaciones de datos abiertos sobre cambio climático
22	¿Cómo debe ser el proceso de apertura de datos?
24	Ejemplos: Plataformas de datos abiertos y cambio climático
26	Datos abiertos sobre cambio climático
27	Datos para abrir y su disponibilidad en Costa Rica
31	Recomendaciones
33	Glosario
34	Referencias
36	Agradecimientos
37	Sobre el autor
38	Sobre la fellowship Estado Abierto

¿Qué es una guía de datos abiertos?

Este documento es una introducción general a qué son los datos abiertos y está enfocado a los datos relacionados al cambio climático. Se espera que este documento sirva como guía para la Dirección de Cambio Climático (DCC) en la implementación de los datos abiertos dentro de sus plataformas web y además sea un incentivo para que esta práctica se difunda entre otras entidades del gobierno.

Este documento es producto del fellowship que brindaron la Iniciativa Latinoamericana por los Datos Abiertos (ILDA) y Escuela De Datos para Costa Rica. Fue desarrollado en la DCC durante la segunda mitad del año 2018.

¿Qué son los datos abiertos?

Los datos se consideran abiertos si pueden accederse, usarse, modificarse y compartirse libremente para cualquier propósito, estando sujetos a lo sumo a medidas que mantengan su autoría y apertura.

Para esto se requiere que los datos sean legalmente abiertos y técnicamente abiertos. Legalmente abiertos se refiere a que los datos estén en el dominio público o que tengan restricciones mínimas de uso. Técnicamente abiertos significa que los datos se publiquen en un formato que pueda ser leído por máquinas (machine readable) y de preferencia en un formato no propietario (formato abierto).¹

Ejemplos de datos que no son abiertos:

Si hay que llenar formularios de inscripción para descargar datos, estos no se consideran como datos abiertos en su definición más estricta.

Si se cobra por descargar ciertos datos, estos tampoco se consideran datos abiertos.

Las plataformas web donde se den listas de datos disponibles pero que no es posible descargarlos no se consideran como datos abiertos.

¹ Ref. Banco Mundial

¿Qué características deben tener los datos abiertos?

Existen ocho principios que deben tener los datos abiertos:

Los datos deben estar completos.

Deben reflejar la totalidad del tema y en el mayor detalle posible.

Los datos deben venir de la fuente primaria.

Deben obtenerse de la fuente primaria con el más alto nivel de detalle posible.

Los datos deben ser accesibles.

Deben estar disponibles para el mayor rango posible de usuarios y propósitos, sin que el usuario tenga que llenar formulario de registro para descargarlos.

Los datos deben ser oportunos y actualizados.

Deben de hacerse disponibles lo más pronto posible y mantener un ritmo de actualización que garantice su utilidad.

Los datos deben ser procesables por máquinas.

Deben de estar en un formato que permita ser procesados automáticamente.

Los datos no deben ser discriminatorios.

Deben estar disponibles para cualquier persona, sin requerir autenticaciones o registro de usuarios.

Los datos deben estar en un formato no propietario.

Deben de servirse en un formato en el no haya ninguna entidad que tenga control exclusivo sobre el mismo. (Esta guía profundiza sobre los formatos abiertos más adelante).

Los datos deben estar licenciados en forma abierta.

Deben de contar con términos de uso y licenciamiento abierto.¹ (Se detalla más adelante).

Otra característica importante que se menciona en la literatura sobre datos abiertos es que estos deberían de poderse descargar como un mismo conjunto de datos (single download), y también que los datos (filas y columnas) deben estar descritos de manera coherente, con nombres que reflejen de forma clara su contenido, y con metadatos que describan cada categoría de una manera precisa.

1 Ref. Guía de datos abiertos en Colombia

Componente legal de los datos abiertos en Costa Rica

Durante la administración Solís Rivera (2014 - 2018) se aprobó el Plan de Acción de la Alianza para un Gobierno Abierto en el que se establecieron una serie de compromisos relacionados a la apertura de datos en sectores que se identificaron como prioritarios.

Su compromiso número cuatro hace referencia a la apertura de “datos públicos relacionados a cambio climático y su financiamiento” y el Ministerio de Ambiente y Energía (MINAE) figura como el ente responsable de su implementación. Este compromiso comprende una serie de actividades y productos que en su mayoría se han desarrollado dentro del proceso de creación del Sistema Nacional de Métrica de Cambio Climático (SINAMECC) de la DCC y que se espera será inaugurado a inicios del año 2019.

Es importante aclarar que no todos los datos públicos son datos abiertos. Como datos públicos entendemos aquellos datos generados por entes estatales en el cumplimiento de sus funciones. A pesar de ser públicos, dadas sus características, la publicación de varios de estos conjuntos de datos estará regulada y/o prohibida por legislación específica. Si los datos son públicos y han sido publicados en algún medio, pero no cumplen con los 8 puntos mencionados en el apartado anterior no se consideran datos abiertos.

Con relación al último párrafo y más específicamente a la legislación costarricense, el informe Diagnóstico del marco legal e institucional en materia de datos abiertos y acceso a la información e implicaciones para el funcionamiento del SINAMECC (Cabrera Medaglia, 2018) menciona entre sus conclusiones que en el país existe un marco jurídico que permite el derecho de petición de información y el derecho a contar con información ambiental, pero también existe legislación que limita el libre acceso a cierta información pública y que esta legislación podría ser aplicable a ciertos conjuntos de datos relacionados a cambio climático.

Con respecto a estas limitaciones, dicho informe hace mención a la Ley de Información No Divulgada y la Ley de Protección de Datos Personales; pero también dice que “en tesis de principio ello no impide la generación de datos o resultados agregados mientras no se individualice a las personas físicas o jurídicas”. (p6)

Según este informe, y en referencia a lo dispuesto en la Constitución “y tratándose de información que un particular no se encuentre obligado a suministrar según el ordenamiento jurídico vigente, cualquier persona —en el ejercicio de su autonomía de la voluntad y del uso de su consentimiento previo— puede condicionar su uso, divulgación o transferencia de los datos proveídos, por ejemplo, mediante acuerdos de confidencialidad y no uso u otras herramientas similares” (p.6).

Este aspecto de confidencialidad de cierta información deberá tomarse en consideración para la gestión de los datos albergados en SINAMECC ya que esto limitará su divulgación.

¿Cuál es la importancia de los datos abiertos?

Al abrir datos, el gobierno garantiza la transparencia en las acciones que realiza, mejorando de esta manera la confianza de la ciudadanía, a través de diversos mecanismos de rendición de cuentas. Los datos, mediante el uso que le den los distintos grupos de usuarios (investigadores, desarrolladores y otros profesionales) logran tener un valor agregado al generar información que se puede apreciar de manera descriptiva o incluso generando información oculta que se puede revelar mediante el análisis estadístico. Un ejemplo de esto son las aplicaciones que se han desarrollado utilizando datos abiertos y que a la vez visualizan relaciones entre variables.

Los datos abiertos también ayudan a realizar una gestión con decisiones basadas en datos. Al tener que recopilar, organizar y distribuir los datos, las distintas instituciones involucradas deben de analizarlos tanto de forma retrospectiva como actual, haciendo posible la comparación y evaluación del efecto de distintas acciones realizadas. Los datos abiertos son una herramienta para la planificación, gestión y evaluación de políticas públicas y cualquier otra disposición de las autoridades.

En el caso de los datos relacionados al cambio climático, se espera que al abrirlos se influya en promover un diálogo entre los diferentes generadores de datos, tomadores de decisiones y la sociedad civil, para así también guiar en la identificación de vacíos de datos y conjuntos de datos importantes que deben de generarse y abrirse. Esta comunicación y participación activa entre distintas entidades también servirá para incentivar un mayor uso de los datos, creación de casos de uso que sean de interés para la sociedad y generar el valor agregado de los mismos con su utilización.

¿Qué son los estándares de datos abiertos?

Los estándares de datos abiertos son documentos que establecen requerimientos y especificaciones que deben de tener los conjuntos de datos para hacerse disponibles al público. Un objetivo importante de estos estándares es —a través de sentar bases comunes—, hacen compatibles los distintos conjuntos de datos y facilitar su identificación y análisis.

Comúnmente los estándares describen los datos de un tema en específico. Por ejemplo, existen estándares para datos de contrataciones públicas o para permisos de construcción, y cada uno posee definiciones y características específicas a su campo.

Tipos de estándares de datos

Estándares esquemáticos

Estos definen la estructura de los datos que se van a publicar. Incluyen, por ejemplo, nombres, descripciones, y tipos de datos de las columnas de las tablas. También pueden incluir relaciones entre distintas tablas.

Estándares semánticos

En este caso el estándar define la terminología y/o lenguaje de los datos que serán publicados. Un ejemplo de este tipo de estándar es el Data Catalog Vocabulary (DCAT), el cual es un vocabulario que tiene definiciones muy claras del significado de cada uno de sus componentes.

Estándares atómicos

Estos definen cómo se deben representar los elementos básicos de los datos en los casos en que haya confusión. Los estándares atómicos pueden definir valores específicos de los datos (cómo se deben de escribir las fechas, por ejemplo), así como también combinaciones de valores. Como muestra, según el estándar ISO8601, la fecha del 18 de octubre del 2018 a las 2 pm debería escribirse como: “2018-10-18T14:00”¹.

En muchos casos los estándares para datos abiertos combinan distintos elementos de los tres tipos mencionados. Por ejemplo, el Open Data Standards Directory menciona que el Estándar de Permisos de Construcción (BLDS por sus siglas en inglés) recomienda publicar dentro del conjunto de datos un elemento llamado PermitClassMapped, este elemento debería de contener uno de dos valores: Residential o Non-Residential. Más adelante en el documento del estándar se define claramente qué significa cada uno de esos valores.

1 Ref. Open Data Standards Directory

Actualmente el World Resources Institute (WRI) está trabajando en un borrador de una guía para datos abiertos de cambio climático llamada Open Up Guide: Using Open Data to Advance Climate Action. El documento pretende servir como una guía que ayude a los distintos actores a priorizar y publicar los tipos de datos más importantes de los contextos domésticos. El documento se elaboró tomando en cuenta datos relacionados a riesgos, impactos, mitigación y medidas de adaptación al cambio climático, y presenta recomendaciones de cuales son los principales datos que deben abrirse en cada una de esas categorías.

¿Qué son las licencias abiertas y por qué son importantes?

Las licencias son herramientas legales que se utilizan para explicar de manera muy clara las condiciones en que se pueden utilizar los datos. Normalmente estas licencias se muestran en una versión resumida con los conceptos principales de la misma, y una versión más extensa donde se da la justificación legal de qué se puede hacer con los datos.

Hay licencias estándar que pueden reutilizarse así como también licencias específicas que ciertos países u organizaciones han desarrollado para casos especiales. Ejemplos de las primeras son la Creative Commons (cc) y la Open Database License (ODbL); y de las segundas son los términos de uso de los datos del Banco Mundial y la licencia abierta del gobierno de Francia.

En el ámbito general de los datos abiertos, las licencias más utilizadas son las Creative Commons en sus versiones de cc-by y cc0. La licencia cc-by se refiere a que el usuario de los datos (beneficiario de la licencia) podrá copiar, distribuir, exhibir, representar y hacer derivados de los mismos siempre y cuando reconozca y cite la fuente de los datos de la forma especificada por el autor. La licencia cc0 o dedicación a dominio público se refiere a que el autor de los datos renuncia a sus derechos de autor, pasando los datos a formar parte del dominio público, permitiendo así que el usuario pueda usarlos libremente. Además de estas, Creative Commons tiene otras variantes que pueden ajustarse a distintas necesidades de los autores de los datos. [Ref. Wikipedia - Creative Commons](#)

¿Qué son los formatos abiertos de datos?

Los datos digitales pueden presentarse en distintos formatos que en muchos casos podemos ver según la terminación de los archivos que los guardan. Por ejemplo, en un archivo llamado informe.doc su terminación .doc nos indica que es un documento en el formato antiguo del programa Microsoft Word. Algunos de estos formatos fueron creados por empresas que limitan su utilización a únicamente los programas creados por ellos mismos, mientras otros de estos formatos fueron creados en estándares abiertos que facilitan su utilización en una mayor variedad de programas.

Una selección de los formatos abiertos más comúnmente utilizados para documentos de texto, tablas e imágenes es:

Documentos de texto

pdf - Acrónimo de portable document file. Originalmente creado por la empresa Adobe como un formato propietario (cerrado) pero desde la versión 2.0 está estandarizado bajo ISO. Es importante mencionar que un documento PDF puede estar compuesto tanto por imágenes escaneadas de documentos físicos como también por documentos electrónicos. La diferencia entre estos dos es que cuando el PDF se compone de imágenes escaneadas, los datos no pueden extraerse tan fácilmente como cuando el pdf se crea directamente, por ejemplo, desde un documento de Word. En el primer caso sería necesario utilizar un software de reconocimiento de caracteres (OCR) para tratar de recuperar parte del texto, en el segundo caso, los datos podrían extraerse automáticamente a través de seleccionar el texto, copiar y pegar.

docx - Formato parte del Office Open XML para documentos de Word desarrollado por Microsoft. Está estandarizado bajo ISO/IEC.

txt - Documento de texto codificado bajo algún código no propietario (como por ejemplo ASCII o UTF-8).

xlsx - Documento de hoja de cálculo parte del Office Open XML para documentos de Excel desarrollado por Microsoft. Está estandarizado bajo ISO/IEC.

Tablas y hojas de cálculo

json - Formato de texto estructurado, acrónimo de JavaScript Object Notation muy popular para el intercambio de datos en aplicaciones web. El formato está estandarizado bajo RFC.

csv - Documento de texto para tablas. En este formato las entradas de cada columna se separan por comas y el texto está codificado en algún código no propietario. Este es el formato abierto preferido para compartir información ordenada como tablas.

Imágenes

GIF - Formato originalmente creado por la empresa CompuServe y abierto desde que venció su patente en 2004.

JPG2000 - Formato para imágenes de mapas de bits (raster) estandarizado bajo ISO/IEC.

PNG - Formato para imágenes de mapas de bits (raster) estandarizado bajo ISO/IEC.

SVG - Formato para imágenes vectoriales estandarizado bajo W3C₁

Para conocer más +

Distintas recomendaciones importantes sobre los formatos a utilizar, nomenclatura de archivos y demás características de los datos abiertos se puede encontrar en el documento **Guía para la publicación de datos en formatos abiertos** de la Secretaría de Modernización del Gobierno de Argentina.

1 Ref. Wikipedia

¿Cómo se evalúan los portales de datos abiertos?

Existen diversas metodologías y maneras de evaluar un portal de datos abiertos. A continuación compartimos algunas de las más utilizadas.

Las 5 estrellas de los datos

Una metodología muy utilizada de cinco escalafones que evalúa y propone diferentes medidas que facilitan el acceso, distribución y análisis de los datos abiertos. La meta de las instituciones que publican datos abiertos es llegar a tener las cinco estrellas, pero en muchos casos, y dependiendo de las capacidades técnicas de las instituciones, esta meta se desarrollará través de un proceso que puede tomar tiempo.

Según este sistema, portales de datos abiertos se clasifican de la siguiente manera:

- * Datos publicados en la web en cualquier formato y utilizando una licencia abierta.
- * * Datos publicados como datos estructurados (tabla de Excel, por ejemplo) en vez de archivos pdfs con tablas o documentos escaneados.
- * * * Datos publicados en formatos no propietarios (csv en vez de Excel, por ejemplo).
- * * * * Se utilizan identificadores de recursos uniformes (URIs - Uniform Resource Identifier) para los conjuntos de datos. De esta forma los usuarios pueden utilizar una ruta específica para acceder el set de datos requerido.
- * * * * * Los datos están enlazados a otros datos, dándoles de esta manera mayor contexto.

Para conocer más +

El sitio de 5* Datos Abiertos ofrece un resumen de costos y beneficios de cada una de estas estrellas.

Tim Berners-Lee del World Wide Web Consortium tiene una presentación en YouTube sobre las ventajas de las cinco estrellas de datos abiertos. El video se puede ver en el enlace: “Open, Linked Data for a Global Community”.

Open Data Certificate

Otro método para evaluar la calidad de los portales de datos abiertos es el open data certificate creado por el Open Data Institute. Este emplea una herramienta gratuita en línea que permite certificar portales de datos abiertos tomando como base el sistema de las cinco estrellas y otros estándares existentes. La herramienta da una calificación basada en cuatro insignias:

- Bronce** Datos con licencia abierta, disponibles sin restricciones, accesibles y legalmente reutilizables.
- Plata** Además de lo anterior, los datos también están documentados en un formato legible por computadora (machine readable), confiables y ofrecen asistencia continua del creador de los datos a través de un canal de comunicación dedicado.
- Oro** Los datos están publicados en un formato abierto legible por computadora, son actualizados regularmente, ofrece mayor documentación de soporte y contienen declaración de derechos legible por computadora.
- Platino** Satisface los requerimientos de las insignias anteriores. Y además los datos tienen documentación de su procedencia legible por computadora, utiliza identificadores únicos y los creadores de los datos tienen un equipo de comunicación ofreciendo asistencia.

Evaluaciones de datos abiertos sobre cambio climático

Cabe resaltar que la organización brasileña IMAFLORA, elaboró un documento donde evaluó utilizando sus propios criterios la apertura de datos relacionados a cambio climático en Brasil. Los criterios tomados en cuenta fueron: disponibilidad en línea, metadatos, actualización, desagregación, legible por computadora (machine readable), en formatos no propietarios, con la opción de descarga única, gratuidad, licencia abierta y acceso directo en línea.

El documento y la metodología utilizada pueden accederse en este **enlace**.

¿Cómo debe ser el proceso de apertura de datos?

La Iniciativa Global para la Transparencia Fiscal elaboró un tutorial que sirve de guía general para esbozar el proceso (a pesar de no tratar directamente sobre datos sobre cambio climático). Este documento concibe un proceso que está conformado por siete pasos:

1 Escuchar la demanda

Un dato tiene valor en la medida en que se utiliza. Es importante consultarle a los distintos actores de la sociedad civil, academia, sector privado, sector público y potenciales usuarios:

- ¿Qué tipo de datos despiertan su interés?
- ¿Quiénes podrían ser las personas/organizaciones usuarias?
- ¿Con qué frecuencia serán utilizados los datos?
- ¿En qué formato les sería más sencillo utilizarlos?
- ¿Cuánto acceso y qué habilidades digitales tienen las personas/organizaciones identificadas?
- ¿Qué objetivos tienen estas personas/organizaciones en relación con los datos de cambio climático?
- ¿Cuánto saben sobre el cambio climático?
- ¿Podemos identificar algún patrón sociodemográfico en las personas/organizaciones usuarias?

2 Identificar fuente y formato de datos

Trabajar en su apertura implica análisis, procesamiento y creación de nuevos procesos para garantizar la sostenibilidad.

3 Identificar fuente y formato de datos

Es importante revisar la calidad de la información que se publica. Con datos sucios, su consulta y procesamiento se convierte en una tarea complicada.

4 Estándares de referencia

La apertura de datos debe ir en concordancia con estándares, ya sean nacionales o internacionales. Esto permitirá que los datos puedan ser comparables y compatibles con otros contextos.

5 Sostenibilidad de la información

La apertura de información supone la ejecución de un proceso constante que considera la realización de diferentes tareas y actividades.

6 Licencias de uso

Las entidades generadoras de datos deben estipular las licencias de uso que atenderán el aspecto ético y legal de la información.

7 Promover el uso de la información

- Garantizar su difusión y acceso mediante estrategias de comunicación y acciones que faciliten la interacción, vinculación y participación de la sociedad en el uso de los datos publicado.
- Desarrollar estrategias de consulta y difusión de la información que permitan compartir experiencias de éxito y desarrollo de capacidades

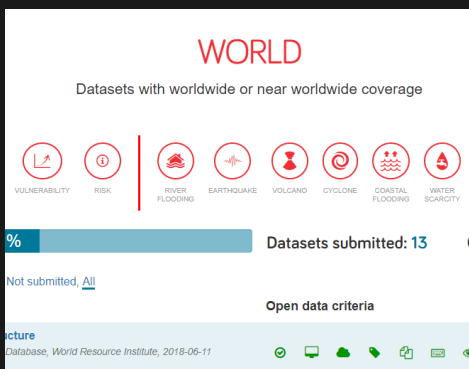
Ejemplos: Plataformas de datos abiertos y cambio climático



Inventario Nacional de Gases de Efecto Invernadero de Argentina

<https://inventariogei.ambiente.gob.ar>

Portal con documentos de los informes y herramientas para crear reportes gráficos interactivos con la opción de descarga de datos de los inventarios de GEI que se han realizado en Argentina.



Open Data for Resilience Index

Índice de Datos Abiertos para la Resiliencia

<https://index.opendri.org>

Plataforma global de datos relacionados a la gestión del riesgo a desastres. Evalúa la apertura de datos por país y permite contribuir con conjuntos de datos estandarizados. Costa Rica no tiene ningún conjunto de datos.

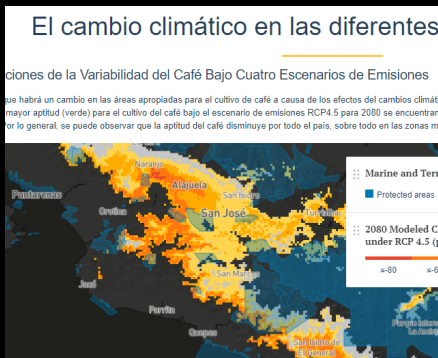


Open Data for Resilience Initiative

Iniciativa de datos abiertos para la resiliencia

<https://opendri.org>

Iniciativa de datos abiertos para la resiliencia. Plataforma global que produce el índice anterior. Se enfoca en tres pilares: compartir, recolectar y reutilizar datos. Muestra los esfuerzos por país en cada uno de estos pilares.



Partnership for Resilience and Preparedness

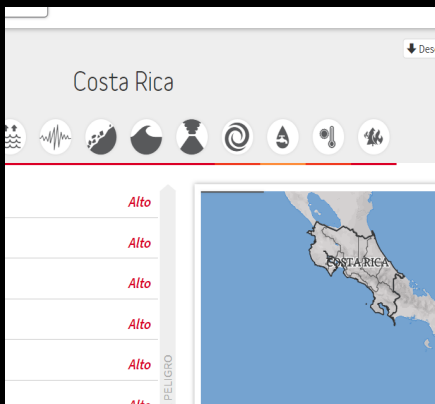
Alianza para la Resiliencia y la Preparación

<https://prepdata.org>

Como establece su lema: “Visualizando datos para crear resiliencia climática”, ésta plataforma cuenta con herramientas para explorar datos abiertos, crear visualizaciones y buscar fuentes de datos relacionados a cambio climático.

Ver el ejemplo: "El futuro del café en Costa Rica":

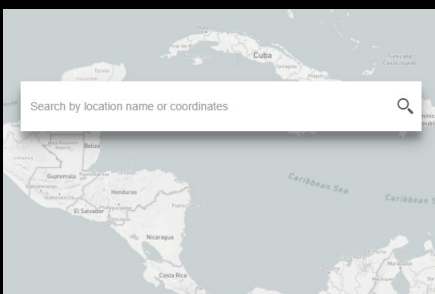
<https://prepdata.org/dashboards/costa-rica-example>



Think Hazard!

<http://thinkhazard.org>

Plataforma para identificar peligros y entender cómo reducir su impacto. Herramienta con mapas y distintos peligros para (casi) todos los países del mundo. Para el caso de Costa Rica la información detalla hasta nivel cantonal. No todos los conjuntos de datos son abiertos, algunos únicamente son para desplegarse como mapas pero no se pueden descargar los datos crudos.



Hazards, Exposures and Vulnerabilities

Data Exploration Tool

<http://assess-risk.info>

Portal de datos sobre peligro, exposición y vulnerabilidad. Actualmente solo tiene datos para algunos países y no está Costa Rica.

Para el caso de Costa Rica, y a pesar de que no son datos estrictamente relacionados a cambio climático, cabe resaltar los esfuerzos que se han realizado con las plataformas de datos abiertos del Gobierno de Costa Rica y Poder Judicial, además de la infraestructura nacional de datos espaciales llamada Sistema Nacional de Información Territorial (SNIT).

Datos abiertos sobre cambio climático

Datos para abrir y su disponibilidad en Costa Rica

Esta es una lista general elaborada tomando como base las recomendaciones del borrador Open Up Guide: Using Open Data to Advance Climate Action del WRI, la comunicación que se ha tenido con el equipo de la DCC y consultas hechas a distintos actores generadores de datos relacionados al cambio climático.

Los conjuntos de datos listados a continuación están agrupados siguiendo las categorías del Open Up Guide y se ajustan a los nombres de los conjuntos de datos existentes en Costa Rica.

Datos sobre impactos y amenazas naturales

Conjunto de datos	Entidad encargada de su generación	Observaciones
Datos meteorológicos (precipitación, temperaturas, etc.)	Instituto Meteorológico Nacional (IMN), Instituto Costarricense de Electricidad (ICE)	Aunque actualmente existe accesibilidad a los resúmenes mensuales del IMN y a los del ICE según acuerdo de confidencialidad, sería importante tener también acceso abierto no solo a los registros mensuales si no también a los diarios de al menos las temperaturas máximas, mínimas y precipitación.
Caudales de ríos	Acueductos y Alcantarillados (AyA), Instituto Costarricense de Electricidad (ICE), Dirección de Agua (DA)	Actualmente existen distintas entidades gubernamentales con estaciones automáticas para mediciones de caudal en varias cuencas del país. Sería importante tener acceso abierto a al menos el listado de esas cuencas aforadas.
Historial de eventos hidrometeorológicos extremos y los daños causados	Comisión Nacional de Prevención de Riesgos y Atención de Emergencias (CNE).	El CNE cuenta con un documento en pdf con el historial de desastres naturales del país y breves reseña de las comunidades afectadas. Como parte del fellowship de ILDA y Escuela de Datos se extrajo la tabla del pdf y se analizaron varios de estos desastres, especialmente los relacionados a eventos hidrometeorológicos.

Datos sobre uso del suelo, cambio de uso y forestería

Conjunto de datos	Entidad encargada de su generación	Observaciones
Cobertura del suelo	Ministerio de Ambiente y Energía (MINAE). Ministerio de Ganadería (MAG).	Los datos del último inventario nacional forestal (2015) ya están abiertos en el SNIT. También en el SNIT se pueden acceder los datos históricos y actuales sobre la extensión del cultivo de la piña. El ICAFE ha trabajado en las estadísticas de las extensiones y productividad del café.
Historial de fincas con pago por servicios ambientales	Fondo Nacional de Financiamiento Forestal (FONAFIFO)	Número, extensión y demás estadísticas de las fincas participantes en el programa de pago por servicios ambientales.

Datos sobre agricultura

Conjunto de datos	Entidad encargada de su generación	Observaciones
Nivel de aguas subterráneas	Dirección de Agua del Ministerio de Ambiente y Energía.	La Dirección de Agua actualmente tiene el SINIGIRH, un sistema que en tiempo real registra los datos del nivel de los acuíferos en ciertos lugares del país. Aunque el acceso al sistema actualmente es restringido, sería importante tener acceso abierto a los registros históricos.

Datos sobre vulnerabilidad climática

Conjunto de datos	Entidad encargada de su generación	Observaciones
Acueductos comunitarios (ASADAS), municipales y administrados por el AyA	Acueductos y Alcantarillados (AyA)	El AyA cuenta con la lista y ubicación de las Asociaciones administradoras de los Sistemas de Acueductos y Alcantarillados comunales (ASADAS).
Acueductos que cuentan con planes de seguridad del agua	Acueductos y Alcantarillados (AyA)	Historial y estadísticas de acueductos comunitarios, municipales y administrados por el AyA que cuentan con planes de seguridad del agua implementados y en proceso de implementación.
Comunidades con planes de emergencia ante eventos climáticos extremos, además de comités municipales y comunales de emergencia	Comisión Nacional de Prevención de Riesgos y Atención de Emergencias (CNE).	
Comunidades, organizaciones y empresas galardonadas por el programa Bandera Azul Ecológica	Programa Bandera Azul Ecológica del AyA.	El programa tiene varias categorías relacionadas a mitigación y adaptación al cambio climático y en su página web se encuentran las listas de organizaciones, empresas y comunidades galardonadas a través de los años.
Datos demográficos	Instituto Nacional de Estadística y Censos (INEC)	Los datos agregados según diferentes variables de los distintos censos realizados por el INEC se pueden consultar y descargar desde su sitio web.

Datos sobre emisiones

Conjunto de datos	Entidad encargada de su generación	Observaciones
Inventario nacional de gases de efecto invernadero	Instituto Meteorológico Nacional (IMN)	Los datos resumen están disponibles actualmente como tablas en documentos pdf.

Datos sobre electricidad

Conjunto de datos	Entidad encargada de su generación	Observaciones
Estadísticas relacionadas a generación eléctrica	Dirección de Energía del Ministerio de Ambiente y Energía (MINAE), Instituto Costarricense de Electricidad (ICE)	El ICE publica informes mensuales en formato pdf con los resúmenes del balance eléctrico por fuente.

Recomendaciones

- + Es de suma importancia que las distintas entidades públicas que generarán datos que se incorporarán al SINAMECC establezcan en los contratos y Términos de Referencia -especialmente cuando estos datos se generan a través de consultorías- que los datos producidos por el proyecto serán publicados como datos abiertos y son datos públicos por lo cual no pueden ser sometidos a tratamientos privados.

- + Existen otros conjuntos de datos prioritarios que se pueden incluir al proceso de apertura, pero que se desconoce su existencia. Entre ellos se pueden mencionar los relacionados al transporte público (posiblemente existan en el Ministerio de Transportes y/o en el Laboratorio Nacional de Materiales y Modelos Estructurales - LanammeUCR), manejo de desechos (en el Informe sobre el Estado de la Nación se hacen referencia a distintas fuentes con estadísticas como el ICE, Ministerio de Salud, distintas municipalidades y las compañías EBI y WWP), datos sobre producción agrícola e inventarios de ganado vacuno (posiblemente en el Ministerio de Agricultura y Ganadería MAG) y datos financieros climáticos (posiblemente la Dirección de Cambio Climático tenga parte de esta información). Para ahondar en los detalles de qué son estos datos, revisar la guía producida por el WRI.

- + Muchos de los datos generados en los Informes del Estado de la Nación son muy relevantes para el análisis de la adaptación y mitigación al cambio climático, por lo que sería un gran aporte si se logra proponer que a futuro los datos publicados en estos informes se compartan como datos abiertos.

Glosario

API Acrónimo de application programming interface. Son un conjunto de herramientas de programación que facilitan la comunicación entre softwares. A través de un API se puede acceder directamente desde un programa de análisis de datos, app o página web los datos de una plataforma de datos abiertos.

CKAN Acrónimo de Comprehensive Knowledge Archive Network. Es una plataforma web de código abierto que se utiliza para catalogar, guardar y distribuir datos abiertos. A pesar de que existen otras plataformas que cumplen esa misma función, CKAN es muy popular entre catálogos de datos abiertos gubernamentales como los de los gobiernos del Reino Unido, Estados Unidos y Australia. La plataforma de datos abiertos del Poder Judicial de Costa Rica utiliza CKAN.

Metadatos Los metadatos o "datos de los datos" es la información que en términos generales nos indica de donde viene un conjunto de datos y sus características. Existen esquemas preestablecidos que indican qué información debe incluirse en los metadatos de cierto tipo de datos. Por ejemplo para datos geográficos existe un esquema estandarizado (ISO 19115:2003) que nos indica qué información debemos adjuntar al conjunto de datos para garantizar su usabilidad. El artículo de Wikipedia sobre metadatos da una guía sobre qué estándares y esquemas existen para distintos tipos de datos.

Referencias

5 estrellas de Datos Abiertos. (s. f.). Recuperado 27 de diciembre de 2018, de <http://5stardata.info/es/>

Cabrera Medaglia, J. (2018). Diagnóstico del marco legal e institucional en materia de datos abiertos y acceso a la información e implicaciones para el funcionamiento del SINAMECC. Dirección de Cambio Climático del Ministerio de Ambiente y Energía y Fundecooperación en el marco del proyecto ICAT.

Creative Commons. (s. f.). Recuperado 27 de diciembre de 2018, de <https://creativecommons.org/>

Definición de Conocimiento Abierto - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge. (s. f.). Recuperado 27 de diciembre de 2018, de <https://opendefinition.org/od/2.1/es/>

Gobierno de la República de Colombia. (2016, junio). Guía de Datos Abiertos en Colombia. Recuperado de http://estrategia.gobiernoenlinea.gov.co/623/articles-8248_Guia_Apertura_Datos.pdf

III Plan de Acción. (2017, junio 12). Recuperado 27 de diciembre de 2018, de <http://gobiernoabierto.go.cr/3pa/>

Open Data Essentials | Data. (2018, diciembre 6). Recuperado 27 de diciembre de 2018, de <http://opendatatoolkit.worldbank.org/en/essentials.html>

Open Data Standards Directory. (s. f.). Recuperado 27 de diciembre de 2018, de <https://datastandards.directory/glossary>

Paquete de Apertura de Datos de la República Argentina. (s. f.). Recuperado 27 de diciembre de 2018, de <https://datosgobar.github.io/paquete-apertura-datos/>

The ODI – Open Data Institute. (s. f.). Recuperado 27 de diciembre de 2018, de <https://theodi.org/>
Wikipedia (s. f.). Recuperado 27 de diciembre de 2018, de https://es.wikipedia.org/wiki/Licencias_Creative_Commons
https://en.wikipedia.org/wiki/List_of_open_formats
https://en.wikipedia.org/wiki/Metadata_standard

Agradecimientos



Iniciativa Latinoamericana para los Datos Abiertos



Sobre el autor

Guillermo Durán



Ingeniero forestal y geógrafo, fellow en cambio climático por la Iniciativa Latinoamericana por los Datos Abiertos y Escuela de Datos. Tiene experiencia en el uso de diferentes tecnologías, como programación en R (tidyverse, shiny y data.table), PostfreSQL, ArcGIS y QGIS. Su trabajo en biogeografía lo ha llevado a diseñar áreas protegidas en Panamá, georeferenciar los datos de los museos de Historial Natural o trabajar en las distribuciones de ciertas especies en futuros climáticos usando Machine Learning. Guillermo estudió ingeniería forestal en Instituto Tecnológico de Costa Rica y luego obtuvo una maestría en Geografía en la Universidad Estatal de San Francisco, California. Forma parte del Centro de Investigaciones Geofísicas de la Univesidad de Costa Rica con el análisis y visualización de modelos climáticos.

Sobre la fellowship Estado Abierto

La Fellowship EstadoAbierto es una oportunidad para entusiastas de la alfabetización y uso de datos desde el sector público promovida por Escuela de Datos y la Iniciativa Latinoamericana para los Datos Abiertos (ILDA). Las fellowships son posiciones cortas para personas apasionadas por los datos, con talento y creatividad para mejorar la liberación y el uso de datos.

Con este programa se busca aumentar la conciencia sobre el uso de datos en el sector público y construir comunidades que, trabajando juntas, usen sus habilidades para la alfabetización de datos. La fellowship EstadoAbierto apoya talleres, eventos de comunidad y contenidos educativos, además de generar proyectos de mayor alcance con las organizaciones aliadas.

En su primera edición, esta Fellowship se enfocó en dos áreas: las Contrataciones Públicas Abiertas y el Cambio Climático y produjo guías como ésta para cada tema. Este programa fue apoyado por el fondo de transparencia del Banco Interamericano de Desarrollo (BID) y contó con la colaboración de la Fundación Avina.

fellowship
ESTADO
ABIERTO
2018