

DATA MANAGEMENT PLAN

1. Data Summary

¿Cuál es el propósito de la recolección/generación de datos y su relación con los objetivos del proyecto?

What is the purpose of the data collection/generation and its relation to the objectives of the project?

Justificar la creencia de que en las grandes ciudades con un gran volumen de tráfico existe a una menor calidad del aire que se respira. Se realiza el estudio en paralelo de la ciudad de Madrid y en la de Barcelona.

¿Qué tipo y formatos de datos se generarán/recolectarán?

What types and formats of data will the project generate/collect?

Se generarán datos en archivos de extensión .csv (*comma separated values*) compuesto de valores, generalmente en *comma flotante*.

¿Se re-usarán datos ya existentes? ¿Cómo?

Will you re-use any existing data and how?

Se reúsan los datos abiertos de los ayuntamientos de las comunidades autónomas a las que pertenecen las ciudades que van a ser estudiadas.

¿Cuál es el origen de los datos?

What is the origin of the data?

Como se ha dicho anteriormente, los datos históricos serán de origen abierto del Ayuntamiento de la Comunidad. Y estos serán complementados con algunos datos creados por nosotros.

¿Cuál es el tamaño esperado de datos?

What is the expected size of the data?

????????

¿A quién le resultará útil ('data utility')?

To whom might it be useful ('data utility')?

- Para el ciudadano, ya que genera tranquilidad a la hora de respirar en la calle (asmáticos, alérgicos).
- Para estudios científicos sobre el aire.
- Para empresas que quieran producir nuevos servicios/productos: coches menos contaminantes.

2. Fair data

- **Makin data findable, including provisions for metadata**

¿Se pueden encontrar los datos producidos y/o usados en el proyecto con metadatos? ¿Se pueden identificar y localizar mediante los estándares de identificación (¿p.e. identificadores persistentes y únicos como DOI)?

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

Sí, el repositorio de UCrema crea identificadores.

¿Cuáles son los estándares de nomenclatura que se siguen?

What naming conventions do you follow?

Se va a usar el estándar de metadatos *Dublin Core* extendido para la generación de los metadatos.

¿Se proporcionan palabras clave que optimizarán las posibilidades de reutilización?

Will search keywords be provided that optimize possibilities for re-use?

Sí. Se listan a continuación:

- | | |
|----------------------------|-------------------------------|
| - Calidad aire Madrid | - Calidad aire Barcelona |
| - Tráfico carretera Madrid | - Tráfico carretera Barcelona |
| - Contaminación Madrid | - Contaminación Barcelona |

¿Se ofrece un control de versiones claro?

Do you provide clear version numbers?

No, ya que no hay versiones.

¿Qué metadatos van a generarse? En el caso de que no exista ningún estándar de metadatos en la disciplina, por favor, describe qué tipos de metadatos son creados y cómo.

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Como se ha dicho anteriormente, se crearán metadatos de los conjuntos de datos que se listan a continuación, siguiendo el estándar de *Dublin Core* extendido:

- Metadatos sobre el proyecto
- Metadatos sobre los datos de la calidad del aire de Madrid
- Metadatos sobre los datos de la calidad del aire de Barcelona
- Metadatos sobre los datos del tráfico de Madrid
- Metadatos sobre los datos del tráfico de Barcelona

- **Makin data openly accessible**

¿Qué datos producidos y/o usados en el proyecto serán publicados abiertamente? Si ciertos conjuntos de datos no se pueden compartir (o son repartidos bajo ciertas restricciones), explicar por qué, claramente separando las razones legales y contractuales provenientes de restricciones voluntarias.

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Todos los datos serán accesibles, tanto los que han sido utilizados, puesto que dichos datos ya son *Open Data*, como los generados en este proyecto.

¿Cómo se harán los datos accesibles (p.e. mediante repositorio)?

How will the data be made accessible (e.g. by deposition in a repository)?

Como el proyecto será llevado a cabo junto con la Universidad de Cantabria, los datos y el proyecto serán subidos al repositorio UCrea.

¿Qué métodos o programas informáticos son necesarios para acceder a los datos?

What methods or software tools are needed to access the data?

Se puede acceder a través de Internet mediante un navegador en las páginas de datos en abierto del ayuntamiento de la comunidad autónoma que, además, disponen de API, así como la página web de UCrea.

Como los archivos en los que serán guardados serán de formato .csv no hará falta ningún software para poder verlos, cualquier bloc de notas será suficiente. Sin embargo, se recomienda utilizar hojas de cálculo tipo *Excel*, *LibreOffice Calc* o *software* tipo *Pandas* de *Python*, *R*.

Si hay restricciones en el uso, ¿cómo se dará acceso?

If there are restrictions on use, how will access be provided?

No hay.

- **Making data interoperable**

¿Son los datos producidos en el proyecto interoperables, esto es, se permite el intercambio y reciclado de datos entre investigadores, instituciones, organizaciones, países etc. (i.e. agregar formatos estándar, tanto como es posible complementar con *software* accesible (*open*), y en particular facilitar combinar con diferentes conjuntos de datos de diferentes orígenes)?

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and facilitating re-combinations with different datasets from different origins)?

Sí, ya que como se ha dicho los datos serán guardados en formato .csv y, además, accesibles abiertamente por internet.

¿Qué vocabularios de datos y metadatos, estándares o metodologías se seguirán para hacer los datos interoperables?

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Metadatos: Dublin Core extendido.

Metodologías: Se incluye código escrito en *Python*, con el paquete *Pandas*.

- **Increase data re-use (through clarifying licenses)**

¿Cuál será la licencia de los datos para permitir el mayor reciclado posible?

How will the data be licensed to permit the widest re-use possible?

Common Creative 4.0.

¿Cuándo serán accesibles los datos para reciclado? Sin embargo, hace que la publicación se retrase, problemas con patentes, especificar por qué y cuánto tiempo se hará, con la intención de que los datos deberán ser accesibles tan pronto como sea posible.

When will the data be made available for re-use? If an-embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Como la publicación del proyecto y los datos se realizarán en el repositorio de UCrea, una vez se publiquen se podrá acceder a ellos.

Por otro lado, el estudio será gestionado por dicho repositorio, y por lo tanto no se sabe cuánto tardará en solucionarse en este momento.

¿Cuánto tiempo se pretende que los datos puedan ser reciclados?

How long is it intended that the data remains re-usable?

El tiempo que el repositorio UCrea mantenga accesible los datos. En este caso, siempre.

¿Se describen los procesos que aseguran la calidad de los datos?

Are data quality assurance processes described?

Sí. Como se ha dicho anteriormente, se incluyen los *scripts* de todos los procesos que se han hecho a la hora de trabajar con los datos.

3. Allocation of resources

¿Cuáles son los costes de hacer los datos de vuestro proyecto FAIR?

What are the costs for making data FAIR in your project?

Los datos ya cumplen los estándares FAIR, por lo tanto, coste cero.

4. Data security

¿Cómo se prevee la seguridad de los datos (incluyendo recuperación de datos, así como seguridad de almacenamiento y transferencia de datos personales)? ¿Están los datos almacenados de forma segura en repositorios certificados para la preservación y curación a largo plazo?

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? Is the data safely stored in certified repositories for long term preservation and curation?

La seguridad de los datos y el estudio dependerá totalmente de los repositorios de los ayuntamientos de Madrid y Barcelona y, además, de la Universidad de Cantabria. No compete a este proyecto hacerse cargo de su seguridad.

5. Ethical aspects

¿Existen problemas legales o éticos que puedan causar problemas a la hora de compartir los datos? Estos pueden ser discutidos en el contexto de la revisión ética. Si es relevante, incluye referencias al capítulo de ética en la DoA

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Los datos no tienen ninguna información sensible que pueda suponer problemas éticos o legales.

¿En los cuestionarios existe información para el consentimiento de preservación a largo plazo, así como compartir los datos personales?

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

No procede. No hay datos personales.