

Descripció de les dades

Introducció

Perquè les dades siguin enteses, interpretades i reutilitzables per altres investigadors, és essencial explicar com es van crear les dades, quin és el context, l'estructura i el contingut de manera comprensible i coherent. A RDR, aquest tipus d'informació s'ha de proporcionar de dues maneres, als camps de metadades i al fitxer README.

 [English version](#)

Els fitxers README

Un fitxer README a RDR, com el propi nom indica: llegeix-me, permet comunicar informació important sobre el dataset, aclareixen possibles preguntes sobre la utilització, creació i/o actualització de les dades. És essencial escriure un bon arxiu README perquè tota la investigació es presenti de forma compacta. El fitxer mínimament ha de contenir el següent:

- Títol del dataset, DOI, informació de contacte
- Mètodes
- Resum de dades i fitxers
- Informació específica de dades
- Condicions de reutilització

Us recomanem que creeu el vostre fitxer README basat en aquestes plantilles.

- *Versió en català*
- *Versió en castellà*
- *Versió en anglès*

El fitxer de text pla està organitzat per blocs i cadascun d'ells conté informacions entre claudàtors i el nivell d'obligatorietat de cada apartat.

Aquest document obligatòriament ha de ser en format .txt, s'ha d'anomenar "README" i es recomana que inclogui la màxima informació possible, encara que la informació no encaixi dins de l'esquema de la plantilla. També és recomanable que la informació s'expressi en anglès, a més de l'idioma original, per afavorir la reusabilitat del dataset.

Atenció!

El fitxer README.txt és obligatori per cada dataset publicat a RDR

Els diccionaris de dades

Un diccionari de dades és un tipus de metadada que enllaça de manera organitzada els noms, definicions i característiques de cadascun dels camps o atributs d'un dataset. El seu objectiu és proveir un llenguatge comú entre l'autor de les dades i els possibles usuaris. A més, ens permeten entendre i interpretar un dataset en proporcionar informació bàsica sobre els camps o variables que conté. Brinden la següent informació:

- Què significa cada camp o variable.
- Quin tipus de dades conte
- Quins valors pot prendre, o si fa servir algun catàleg.
- Si conté informació pública, confidencial o reservada.

El diccionari de dades estan dissenyats per facilitar la comprensió i proveir de sentit, per tant han de documentar l'existència, el significat i l'ús de cada element del dataset.

Les persones responsables de les dades han de mantenir actualitzat el contingut del diccionari de dades, incloses les definicions i els valors.

Codebooks

Un codebook proporciona informació sobre l'estructura, el contingut i la disposició d'un fitxer de dades. Un codebook ben documentat conté informació que pretén ser completa i autoexplicativa per a cada variable d'un fitxer de dades.

Tot i que els llibres de codis varien àmpliament en qualitat i quantitat d'informació proporcionada, un llibre de codis típic inclou:

- Ubicacions de columnes i amplades per a cada variable.
- Definicions de diferents tipus de registre.
- Codis de resposta per a cada variable.
- Codis utilitzats per indicar no resposta i dades que falten.
- Preguntes exactes i patrons de salt utilitzats en una enquesta.
- Altres indicacions del contingut i característiques de cada variable.

El cos d'un codebook descriu el contingut del fitxer de dades.