

Instructie voor het maken van een README.txt bestand

Een README.txt bestand bevat informatie over de dataset die niet tot uitdrukking kan worden gebracht in de metadatavelden. Het document helpt toekomstige gebruikers je dataset beter te begrijpen.

Deze instructie is specifiek bedoeld voor het schrijven van een README.txt bestand voor een dataset die gedeponeerd wordt in DANS Easy. Het bevat aanvullende informatie bij de verplichte metadatavelden gehanteerd door DANS Easy (Dublin Core metadata standaard) in de vorm van algemene tips voor het schrijven van en mogelijke inhoud van je README.txt bestand. Het betreft alle mogelijke informatie die je kan opnemen in een README.txt bestand. Maak een keuze voor die informatie die voor jouw dataset relevant zijn. Denk daarbij aan wat iemand anders zou willen en moeten weten over deze dataset om hem te kunnen hergebruiken. Probeer het kort te houden. Maak het niet langer dan 1-2 A4 pagina's.

Praktische tips

- Gebruik zo veel mogelijk een gestandaardiseerde woordenschat en respecteer de conventies die gangbaar zijn binnen je vakgebied

Aanbevolen inhoud

1. Titel van een dataset

2. Context en herkomst van dataset – aanvullende informatie bij de beschrijving

- **Aanleiding: beschrijf kort het onderzoekproject, onderzoeksvragen, etc.** (vooral belangrijk als je geen publicaties vermeldt gebaseerd op deze dataset, of wanneer je dataset meer data bevat dan alleen de data geanalyseerd in gerelateerde publicaties). Probeer het kort te houden (max. 200 woorden)

3. Informatie over dataset

- Beschrijf de opbouw van een dataset: mappenstructuur, uitleg over hoe de data zijn verdeeld, volgorde van bestanden, relaties tussen mappen en bestanden
- Bestandsnamen – uitleg over de naamconventie (bijvoorbeeld uitleg afkortingen)

4. Methodologie

- Afhankelijk van discipline, kan je denken aan volgende informatie:
 - Het doel waarvoor de data zijn verzameld
 - Instrumenten gebruikt bij het verzamelen van data (survey, vragenlijsten, interviews)
 - Uitleg over de sampling methodologie
 - Keuze deelnemers, beperking
 - Veranderingen in methodiek
 - Benodigde software om de data te openen, te analyseren
 - Andere benodigde technische informatie om de data te hergebruiken
 - Normen en kalibratie
 - Procedures voor de kwaliteitsborging van de data
 - Eventuele beperkingen van de dataset

5. Specifieke informatie over data

- Uitleg over codes en symbolen als er geen documentatie hierover in de databestanden zijn toegevoegd
- Afkortingen binnen bestanden
- Eenheden en afmetingen
- Symbolen van ontbrekende gegevens
- Kleuren

6. Versies van data

- Informatie over methode van dataverwerking
- Veranderingen aangebracht in ruwe data

Eventuele bijlagen: data dictionaries, codebooks, etc.