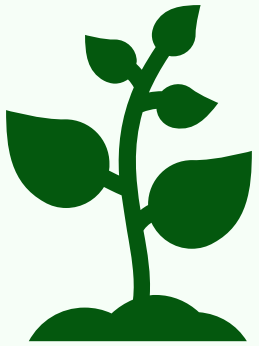
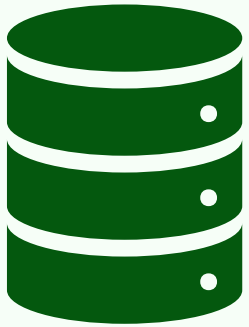
AgriMarket
Agricultural data
at your fingertips

The problem



In agriculture adoption of dataspaces is still in early stages



Sharing data between different actors is difficult when metadata is poorly described

Our solution

Intuituve and modern desing

AI-driven data-analysis

Metadata creation according to Dublin Core and DCAT standards

4 step masterplan

of AI 🤷

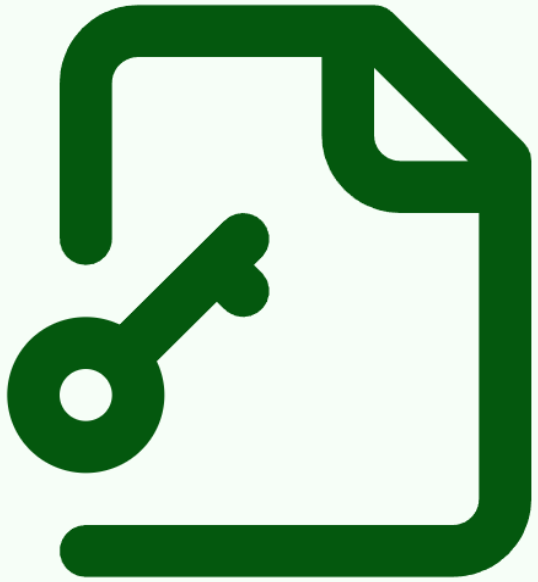
Remember to thank the AI
before it's too late 🤖

1. Unzip

Traverse the directory

```
Directory Structure:
└── Challenge C
    ├── Challenge C links.txt
    └── 3_OpenDroneMap
        ├── odm_texturing
        │   ├── odm_textured_model_geo_material0002_map_Kd.png
        │   ├── odm_textured_model_geo_material0005_map_Kd.png
        │   ├── odm_textured_model_geo_material0008_map_Kd.png
        │   ├── odm_textured_model_geo_material0000_map_Kd.png
        │   ├── odm_textured_model_geo.conf
        │   ├── odm_textured_model_geo.mtl
        │   ├── odm_textured_model.mtl
        │   ├── odm_textured_model_geo.glb
        │   ├── odm_textured_model_geo_material0009_map_Kd.png
        │   ├── odm_textured_model_geo_material0011_map_Kd.png
        │   ├── odm_textured_model_geo_material0013_map_Kd.png
        │   ├── odm_textured_model_geo_material0006_map_Kd.png
        │   ├── odm_textured_model_geo_material0012_map_Kd.png
        │   ├── odm_textured_model_geo_material0010_map_Kd.png
        │   ├── odm_textured_model_geo.obj
        │   ├── odm_textured_model_geo_material0007_map_Kd.png
        │   ├── odm_textured_model_geo_material0001_map_Kd.png
        │   ├── odm_textured_model_geo_material0008_map_Kd.png.aux.xml
        │   ├── odm_textured_model_geo_material0004_map_Kd.png
        │   ├── odm_textured_model_geo_material0014_map_Kd.png
        │   └── odm_textured_model_geo_material0003_map_Kd.png
        ├── odm_report
        │   ├── stats.json
        │   ├── report.pdf
        │   └── shots.geojson
        ├── cameras.json
```

2. Key files



Identify the files with key information

```
Files of Interest: [
  {
    file: "Challenge C/Challenge C links.txt",
    reason: "This text file may contain links that point to releva
  },
  {
    file: "Challenge C/3_OpenDroneMap/cameras.json",
    reason: "As a JSON file, it likely contains structured data ab
  },
  {
    file: "Challenge C/3_OpenDroneMap/odm_report/stats.json",
    reason: "This JSON file probably contains statistical informat
  },
  {
    file: "Challenge C/3_OpenDroneMap/odm_georeferencing/odm_georе
    reason: "This file likely contains valuable metadata about the
  },
  {
    file: "Challenge C/3_OpenDroneMap/odm_orthophoto/odm_orthophot
    reason: "The TIF file is an image file that may hold significa
  }
]
```

3. Summarize

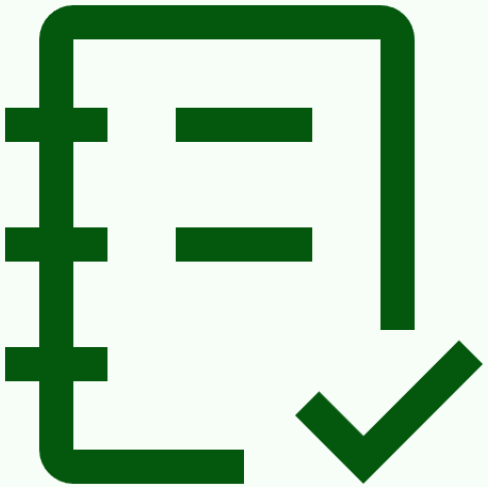
Summarize information from the key files and file names

Summary: The excerpts from the provided files indicate an initiative focused on agricultural data, specifically under the framework of an Agricultural Data Marketplace (Challenge C). Key components include links to valuable resources such as the AGROVOC vocabulary, a multilingual thesaurus that enhances data discoverability in agricultural domains, accessible through the FAO link.

One major aspect of the challenge involves the application of drones and OpenDroneMap (ODM) technology for agricultural data collection and analysis. The cameras used in this context are detailed in JSON format, specifying parameters that adhere to precision requirements vital for applications in crop monitoring and precision agriculture. Drones capture aerial imagery to process various statistics, including feature extraction and reconstruction metrics, which are crucial for creating detailed maps that support agricultural management practices.

Another key data point is the camera data …

4. Standardize



Dublin Core Metadata Initiative®
Making it easier to find information.

Take all the previous information and ask GPT to fill the fields

```
export const dublinCoreSchema = z.object({
  title: z.string().describe("A name given to the resource."),
  creator: z.string().optional().describe("An entity primarily responsible for making the resource."),
  subject: z.string().optional().describe("The topic of the resource."),
  description: z.string().optional().describe("An account of the resource."),
  publisher: z.string().optional().describe("An entity responsible for making the resource available."),
  contributor: z.string().optional().describe("An entity responsible for making contributions to the resource."),
  date: z.string().optional().describe("A point or period of time associated with an event in the lifecycle of the resource."),
  type: z.string().optional().describe("The nature or genre of the resource."),
  format: z.string().optional().describe("The file format, physical medium, or dimensions of the resource."),
  identifier: z.string().optional().describe("An unambiguous reference to the resource within a given context."),
  source: z.string().optional().describe("A related resource from which the described resource is derived."),
  language: z.string().optional().describe("A language of the resource."),
  relation: z.string().optional().describe("A related resource."),
  coverage: z.string().optional().describe("The spatial or temporal topic of the resource, spatial applicability, or jurisdiction."),
  rights: z.string().optional().describe("Information about rights held in and over the resource."),
});

export type DublinCoreMetadata = z.infer<typeof dublinCoreSchema>;
```

```
const completion =
  await openai.beta.chat.completions.parse(
    {
      model: "gpt-4o-mini",
      messages: [
        {
          role: "user",
          content: prompt,
        },
      ],
      response_format:
        zodResponseFormat(
          dublinCoreSchema,
          "dcmi_metadata"
        ),
    }
  );
```

```
Generated DCMI Metadata: {
  title: "Agricultural Data Marketplace Challenge C",
  creator: "Not Available",
  subject: "Drone Technology, Agricultural Data, Remote
  description: "This dataset encompasses drone-collected
ces and technologies pertinent to precision agriculture,
  publisher: "Not Available",
  contributor: "Not Available",
  date: "Not Available",
  type: "Dataset",
  format: "application/json, application/txt, applicatio
  identifier: "Not Available",
  source: "Challenge C Links, OpenDroneMap Outputs, Agri
  language: "English",
  relation: "AGROVOC Vocabulary",
  coverage: "Global, Agricultural Regions",
  rights: "Not Available",
}
```

Prompts 🗣

```
const prompt = `
  You are an AI assistant tasked with analyzing
  the following directory structure to identify
  files that are likely to contain valuable information
  for metadata generation in their first 1000 bytes of ascii.
  Consider factors such as file types, names, and typical content
  relevance. Please list the files of interest along with a brief
  explanation for each selection. List maximum 5 files and make
  sure the file paths are correct and work.

  Directory Structure:
  ${directoryStructure}
`;
```

```
const summarizationPrompt = `
  Based on the following file excerpts, please summarize
  the key information, highlighting agriculture-related data.
  Use AGROVOC terms where appropriate. The summary should be in free-form text.

  File Excerpts:
  ${fileExcerptsText}
`;
```

```
const mainPrompt = `
 Directory Structure:
 ${dirStructure}

 All file extensions:
 ${extensions.join(", ")}

 Summary:
 ${summaryOrFilesKeyInfo}

 Using this information, generate
 Dublin Core Metadata Initiative (DCMI)
 metadata for this dataset. Ensure that
 each metadata element accurately reflects
 the content and context of the files. Use
 MIME types for the dc:format If certain metadata
 elements cannot be determined from the provided
 information, please indicate them as "Not Available."

 dc:identifier: ${id}
`;
```

Website idea 🤔

Upload

Haku

Data itemi jtn muita metadataa

Data itemi jtn muita metadataa

Data itemi jtn muita metadataa

Data itemi jtn muita metadataa

Data itemi jtn muita metadataa

Data itemi jtn muita metadataa

UPLOAD

File upload
Select or drop your file

SUBMIT

Edit

Otsikko

METADATAA

METADATAA

METADATAA

ESIKATSELU JA MUOKKAUS

Tallennus

💎 Live demo 🤞