OnlinePDFLoader

```
class langchain_community.document_loaders.pdf.OnlinePDFLoader(
    file_path: str | PurePath,
    *,
    headers: dict | None = None,
) #
```
[source]

Load online PDF.

Initialize with a file path.

Parameters:

- **file_path** (str | PurePath) – Either a local, S3 or web path to a PDF file.
- **headers** (dict | None) – Headers to use for GET request to download a file from a web path.

Attributes

source	

Methods

__init__ (file_path, *[, headers])	Initialize with a file path.
alazy_load ()	A lazy loader for Documents.
aload ()	Load data into Document objects.

`lazy_load` ()	A lazy loader for Documents.
`load` ()	Load documents.
`load_and_split` ([text_splitter])	Load Documents and split into chunks.

__init__(

  file_path: str | PurePath,

  *,

  headers: dict | None = None,

) #

Initialize with a file path.

Parameters:

- **file_path** (str | PurePath) – Either a local, S3 or web path to a PDF file.
- **headers** (dict | None) – Headers to use for GET request to download a file from a web path.

async alazy_load() → **AsyncIterator[Document]** #

A lazy loader for Documents.

Return type:

  AsyncIterator[Document]

async aload() → **list[Document]** #

Load data into Document objects.

Return type:

list[Document]

lazy_load() → Iterator[Document] #

A lazy loader for Documents.

Return type:

Iterator[Document]

load() → list[Document] #   [source]

Load documents.

Return type:

list[Document]

load_and_split(

  text_splitter: TextSplitter | None = None,

) → list[Document] #

Load Documents and ____ into chunks. Chunks are returned as Documents.

Do not override this method. It should be considered to be deprecated!

Parameters:

text_splitter (Optional[TextSplitter]) – TextSplitter instance to use for splitting documents. Defaults to RecursiveCharacterTextSplitter.

Returns:

List of Documents.

Return type:

list[Document]

Examples using OnlinePDFLoader

- Maritalk