a-parida12 / pdf2dcm

Code | Issues 4 | Pull requests 2 | Discussions | Actions | Projects | Wiki | Security 2 | Insights | Se

Python Package for PDF to DICOM Conversion

a-parida12.github.io/pdf2dcm/pdf2dcm/

MIT license

Code of conduct

20 stars 4 forks 3 watching 4 Branches 8 Tags Activity

Public repository

main | 4 Branches 8 Tags | Go to file t | Go to file | + | Add file | Code | ···

dependabot[bot] build(deps-dev): bump jinja2 from 3.1.3 to 3.1.4 (#63) ✓ ea4a919 · 2 months ago

.github/workflows	ci(semrel): update version (#57)	7 months ago
pdf2dcm	feat(repersonlization): add customizable fields (…	7 months ago
tests	feat(repersonlization): add customizable fields (…	7 months ago
.gitignore	test(base-coverter): test functions of base conve…	2 years ago
.pre-commit-config.yaml	test(repersonalisation): improved check for uid	2 years ago
CHANGELOG.md	0.5.0	7 months ago
CODE_OF_CONDUCT.md	Create CODE_OF_CONDUCT.md	2 years ago
LICENSE	docs(license): update copyright year(s) (#52)	7 months ago
README.md	feat(repersonlization): add customizable fields (…	7 months ago
poetry.lock	build(deps-dev): bump jinja2 from 3.1.3 to 3.1.4 …	2 months ago
pyproject.toml	update version in toml	4 months ago

pdf2dcm

pypi v0.5.0 | python 3.8 | 3.9 | 3.10 | 3.11 | 3.12 | PyPi Velocity 1k | PyPi Downloads 17k | License MIT | codecov 98% | Python Application Testing passing | Test and Release passing

PDF to DICOM Converter

> A python package for PDF to Encapsulated DCM and PDF to DICOM RGB converter

SETUP

Python Package Setup

The python package is available for use on PyPI. It can be setup simply via pip

```
pip install pdf2dcm
```

To the check the setup, simply check the version number of the `pdf2dcm` package by

```
python -c 'import pdf2dcm; print(pdf2dcm.__version__)'
```

Poppler Setup

Poppler is a popular project that is used for the creation of Dicom RGB Secondary Capture. You can check if you already have it installed by calling `pdftoppm -h` in your terminal/cmd. To install poppler these are some of the recommended ways-

Conda

```
conda install -c conda-forge poppler
```

Ubuntu

```
sudo apt-get install poppler-utils
```

MacOS

```
brew install poppler
```

PDF to Encapsulated DCM

Usage

```
from pdf2dcm import Pdf2EncapsDCM

converter = Pdf2EncapsDCM()
converted_dcm = converter.run(path_pdf='tests/test_data/test_file.pdf', path_template_dcm='tests/test_data/CT_small.d
print(converted_dcm)
# [ 'tests/test_data/test_file.dcm' ]
```

Parameters `converter.run` :

- `path_pdf (str)` : path of the pdf that needs to be encapsulated
- `path_template_dcm (str, optional)` : path to template for getting the repersonalisation of data.
- `suffix (str, optional)` : suffix of the dicom files. Defaults to ".dcm".

Returns:

- `List[Path]` : list of path of the stored encapsulated dcm

PDF to RGB Secondary Capture DCM

Usage

```
from pdf2dcm import Pdf2RgbSC

converter = Pdf2RgbSC()
converted_dcm = converter.run(path_pdf='tests/test_data/test_file.pdf', path_template_dcm='tests/test_data/CT_small.d
print(converted_dcm)
# [ 'tests/test_data/test_file_0.dcm', 'tests/test_data/test_file_1.dcm' ]
```

Parameters `converter.run` :

- `path_pdf (str)` : path of the pdf that needs to be converted

- `path_template_dcm (str, optional)` : path to template for getting the repersonalisation of data.
- `suffix (str, optional)` : suffix of the dicom files. Defaults to ".dcm".

Returns:

- `List[Path]` : list of paths of the stored secondary capture dcm

Notes

- The name of the output dicom is same as the name of the input pdf
- If no template is provided no repersonalisation takes place
- It is possible to produce dicoms without a suffix by simply passing `suffix=""` to the `converter.run()`

Repersonalisation

It is the process of copying over data regarding the identity of the encapsualted pdf from a template dicom. Currently, the fields that are repersonalised by default are-

- PatientName
- PatientID
- PatientSex
- StudyInstanceUID
- ~~SeriesInstanceUID~~
- ~~SOPInstanceUID~~

The fields `SeriesInstanceUID` and `SOPInstanceUID` have been removed from the repersonalization by copying as it violates the DICOM standards.

You can set the fields to repersonalize by passing repersonalisation_fields into `Pdf2EncapsDCM()` , or `Pdf2RgbSC()`

Example:

```
fields = [
    "PatientName",
    "PatientID",
    "PatientSex",
    "StudyInstanceUID",
    "AccessionNumber"
]
converter = Pdf2RgbSC(repersonalisation_fields=fields)
```

README Code of conduct MIT license

Releases 8

v0.5.0 Latest
on Jan 31

+ 7 releases

Packages

No packages published
Publish your first package

Contributors 7

Deployments 8

github-pages 7 months ago

+ 7 deployments

Languages

- Python 100.0%