**Presidio Custom Language and PatternRecognizer**

<https://microsoft.github.io/presidio/analyzer/developing_recognizers/>

Create a PatternRecognizer class I put the file on pii\_codex\_services\_analyzers Ex. EsNumberRecognizer

Also, you can define custom PatternRecognizer like we have in recognizers.py in the same Folder.

Create an analyzer class like we have in analisys\_es 🡪 file analysis\_es.py.

Create the logic to load the language and analyzer 

Setup to load the language you consider with install\_spacy\_package("en\_core\_web\_lg")

In this example we are using 2 custom PatternRecognizer **mister\_es** and **es\_number**

In the folder Models, we have all Dataclass we have defined in our application, by default English it is using MSFTPresidioPIIType in file microsoft\_presidio\_pii.py and Spanish PIIType\_ES in common.py. Recommendable create your own class for your language.

Add the custom entities if not exists to pii\_type\_mappings.csv in data\v1

Ex. SPANISH DNI,ES\_NIF,Secure Identifiers,Directly PII,Stand Alone PII,PHI,Identifiable

In the file analysis\_service.py

Create the logic to map Entities Language Dataclass to your Custom Analysis engine



In the endpoint file create the PIIAnalysisService with the appropriate language

