Document

```
id: long
file: string
t2Source: string
t1Source: string
tokenizedSource: string
originalSource: string
license: string
url: string
idiomatic:boolean
```

```
+Document(args*)

+getId(): long

+getSource(): string

+setSource(source: string)

+getT2Source(): string
+getT1Source(): string
+setFile(file:string)
+isIdiomatic(): boolean

+setIdiomatic(idomatic:boolean)
+getRecommendIdiom(): string
+setRecommendIdiom(idiom:string)
```

User

```
+interval: string
```

```
+Main()
+Analyzerepo(): csv
+Analyzerepo(interval:string): csv
+CallBokeh()
+CallSiamese()
```

Bokeh
Pandas

use

use

Elasticsearch database

```
id: int
file: string
startline: int
endline: int
src: string
t2src:string
t1src:string
tokenizedsrc: sring
orgsrc:string
license:string
idomatic: boolean
recommend: string
```

CSV

```
+commit_Id: int
+timestamp_id: int
+idiom_name: string
+sample_id: string
+style: boolean
```

Libraries