

ADD-ON: GALILEO XAI - NER

Il prodotto è in grado di effettuare una NER (Named-entity recognition) su un testo qualsiasi senza un formato predefinito; in particolare nelle categorie: persone e organizzazioni.

Caratteristiche funzionali:

Dato un testo o un file nei formati più comuni (pdf, word, html, txt, open office, etc ..) il prodotto è in grado di esaminare il testo e ricavare l'elenco di tutte le entità presenti, ogni entità è mappata in un file xml di output; nel file xml sono presenti i singoli tag oltre ai tag complessivi, il grado di accuratezza del riconoscimento (sia totale sia sui singoli tag), la/le posizioni dell'entità nel documento.

Ad esempio in un testo semplice come il seguente il prodotto consente di generare un XML sulla falsariga di quello seguente (con testi più complessi alcuni tag sia dell'azienda sia della persona potrebbero essere mancanti o identificati con ranking di accuratezza minore, nel qual caso il ranking dell'entità complessiva potrà essere un derivato dei singoli ranking):

```
il sottoscritto Mario Rossi nato a cesena il 13/3/2020 e residente a Milano in via Cimabue 3
CAP 20124 titolare dell'azienda Verdi spa con CF 12345678901 <persona ranking=100>
<nome ranking=100>Mario</nome> <cognome ranking=100>Rossi</cognome>
<data_nascita ranking=100>13/3/1977</data_nascita> <luogo_nascita
ranking=100>cesena</luogo_nascita> <provincia_nascita
ranking=100>FC</provincia_nascita> <sexo ranking=100>M</sexo> <cf
ranking=100>RSSMRA77C13C573U</cf> <indirizzo_residenza ranking=100> <toponimo
ranking=100>via</toponimo> <desctoponimo ranking=100>Cimabue</desctoponimo>
<civico ranking=100>3</civico> <cap ranking=100>20124</cap> <citta
ranking=100>Milano</citta> <provincia ranking=100>MI</provincia>
</indirizzo_residenza></persona><organizzazione ranking=100> <denominazione
ranking=100>Verdi</denominazione> <tipo ranking=100>spa</tipo> <cf
ranking=100>12345678901</cf>
```

Caratteristiche tecniche del prodotto:

1. è installabile su S.O. open source
2. completamente interoperabile tramite web services (REST)
3. riconoscimento funzionante out of the box: motore pre allenato.