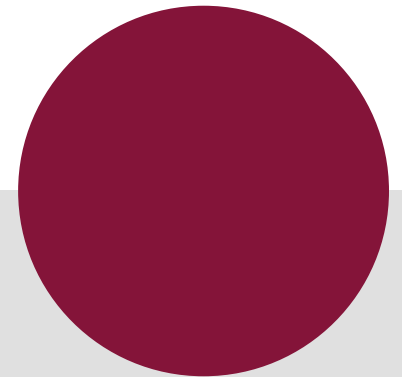




.consulting .solutions .partnership



SAKI 2019 Exercise #2

„Resume NER“



John Loutzenhiser

Lead IT Consultant – Public Sector

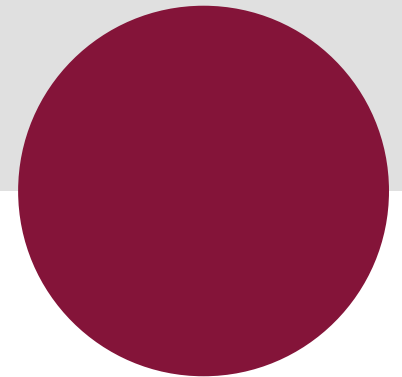
+49 (0) 173 /859 4235

John.loutzenhiser@msg.group

msg systems ag

Robert-Buerkle-Str. 1, 85737 Ismaning
Germany

www.msg.group



.consulting .solutions .partnership

KI in der öffentlichen Verwaltung

- Enorm viel Potential
- Enorm viel Bedarf
- Enorm viel Daten
- Enorm viel Wirkung

NLP in der öffentlicher Verwaltung

- Dokumente
- Entscheidungen

NLP bei der Bundesagentur für Arbeit

- Dokumentenklassifizierung / Automatisierung
 - „Kursnet“ – Klassifikation von Kursbeschreibungen
 - „Stellenmeldeformular“ – unstrukturiert (email) zu strukturiert (Formular)
 - „OPAL“ – Online Kindergeldbeantragung
- Textanalyse bei Stellenangeboten
 - „Mehrfachbeauftragungen“
 - „Skills Mining“
- „Matching“

Resume NER

- „Named Entity Recognition“ – Anspruchsvolle Textanalyse
- Resume Datensatz öffentlich verfügbar
- Datensatz vorannotiert (pre-labeled)
- Wenig Theorie, viel Praxis

Resume NER

- Setup
 - Python (≥ 3.6)
 - https://github.com/jloutz/Resume_NER
 - Project folder with virtenv/venv
 - Jupyter Notebook
- Data Exploration (Notebook)
 - Basic Python, Data exploration
 - No special dependencies
- Spacy NER (Notebook)
 - More Python
 - Train NER Model, Evaluate Model (Evaluation Metrics)
 - Dependencies
 - Pip install – numpy, pandas, spacy
 - Spacy download „en“
 - Pip install flair –no-deps
- Flair NLP (Notebook & Google Colab)
 - Google Account
 - Google drive, Google Colab
 - Train NER using Cloud, Deep Learning & GPU