**Aufgabenstellung:**

Sanchez et al. (s.u.) untersuchen das "framing" von Migration/MigrantInnen in spanischen Parlamentsreden. Wir wollen das, in kleinerem Umfang, für den deutschen Bundestag nachbilden. P2a befasst sich mit Korpuserstellung und Annotation: Erstellen Sie ein Bundestag-Korpus für die Jahre 2015-2017 und konzipieren Sie das Vorhaben manueller Annotation: Werkzeug auswählen, kleine Annotationsrichtlinie formulieren, Annotationen durchführen, Übereinstimmung messen. P2b befasst sich mit autom. Klassifikation: Verwenden Sie ein Verfahren Ihrer Wahl (zB eines, mit dem Sie schon vertraut sind), um anhand des annotierten Korpus das Framing automatisch zu identifizieren, und evaluieren Sie die Performanz. (Für beide Teilprojekte stehen zwei Ko-AutorInnen des Sanchez-Papiers gern zur Konsultation zur Verfügung.)

=> In 2-3 Sätzen die **Grundidee** von Sanchez et. al. erklären   
(neuer Ansatz Stereotype zu identifizieren (subtilere Formen der Stereotypisierung), und zwar mithilfe von Frames)  
**[ Kalie ]**

|  |  |  |
| --- | --- | --- |
| **Schritt** | **Fragen, Erläuterungen** | **Wer?** |
| **Daten wählen** | * Wie unterscheiden sich die Daten (Bundestagsplenarprotokolle) von den Daten („dialogische“ (?) Reden aus dem spanischen Parlament) von Sanchez et. al.? * Inwiefern könnte sich das auf die Ergebnisse auswirken? | **[ Kalie ]** |
| **Preselection & Annotation** | * Kurz Vorgehen beschreiben, evlt. begründen * Annoationsrichtlinien geändert, warum? * Stichwortliste geändert, warum? * Wie ist die Verteilung der Klassen?  (1. Stereotype vs. Nonstereotype; 2. Victim vs. Threat (Kategorien)) | **[ Kalie ]** |
| **Preprocessing** | - Kurz Vorgehen beschreiben, evlt. begründen | **[ Kalie ]** |
| **Vorüberlegungen & Ziele** | * **Sanchez et. al.: RQ3:** *„How effective are classcical machine learning and state-of-the-art transformers models at distinguishing different categories of stereotypes about immigrants with this taxonomy?“* * Unsere konkrete Aufgabe:Support Vector Machine mit Ngrammen und/oder Sentiment, (Tfidf) trainieren  **=>** **Leitfrage:** Wie effektiv und warum?  (Wahrscheinlich läuft hinaus auf ‚Warum so schlechte Performance?‘) | **[ Liane ]** |
| **Modell trainieren** | - Kurz Vorgehen beschreiben | **[ Liane ]** |
| **Evaluation** | **1. Metriken** - Accuracy, , P, R, F1-Score mit Baseline vergleichen (kurz)  **2. PMI**  - Sanchez et. al.:  „*From the most relevant n-grams from the examples labelled as Stereotype, we confirmed that they are not trivially associated with the immigrant group (they are not always biased attributes) and, therefore, we should not rely only on a set of keywords to represent it. Additionally, we confirmed that in the study of stereotypes about immigrants, we have to consider not to define two opposite social groups, since the nonstereotypical texts are neutral phrases talking about immigration in general without stereotyping at all“*  - Zeigt sich das auch bei unseren Ergebnissen (Ngramme mit höchstem PMI)?  **3. Konfusionsmatrix** - Interpretation  **4. Beispiele für richtige/falsche Vorhersagen des Modells** - Warum scheitert das Modell in falschen Beispielen?  **5. Fazit**   * Klassisches Modell vs. komplexeres Transformermodell? * Aufgaben (Stereotype vs. Nonstereotype & Victim vs. Threat) vergleichen * Generelle Probleme (z.B. Frames zu komplex für kurze Sätze?) * Zu wenig Daten * Rhetorik usw. zu komplex für automatische Klassifikation? | **[ Kalie ]**  **[ Liane ]**  **[ Kalie & Liane ]** |

* **Bis Montag Abend:**
  + - fertig annotieren!
* **Bis Dienstag Abend:**
  + - Alles in Tabelle vor Evaluation fertig
    - Liane: Code fertig
* **Mittwoch:**
  + - **Am frühen Mittag discorden (vllt. vor 12???)**
    - Evaluation
* **Donnerstag & Freitag:**
  + - Zusammenfügen & Korrektur usw.