# Al-enabled Language Processing of Adolescents' and Young Adults' Challenges and Aspirations (ALPACA)



Dr. Christina Haag & Dr. David Bürgin







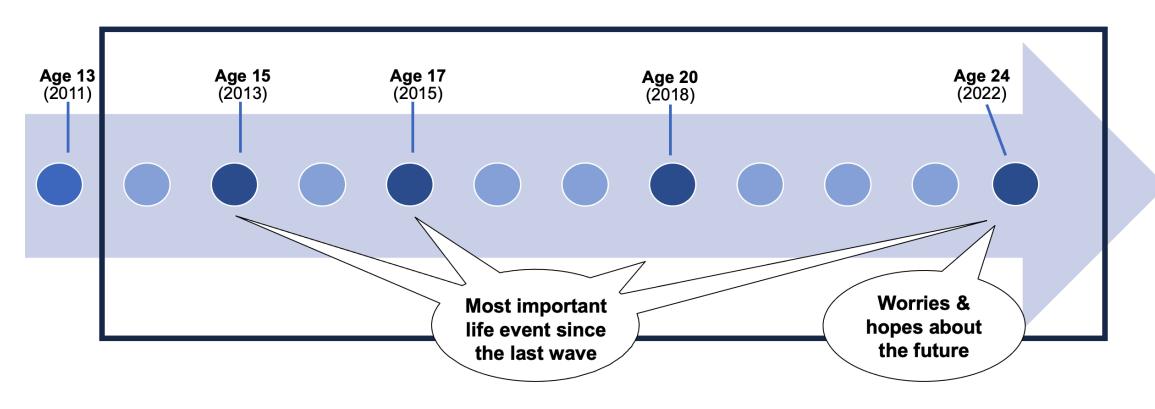


Seed grant from the

**UZH PRC**Population Research Center











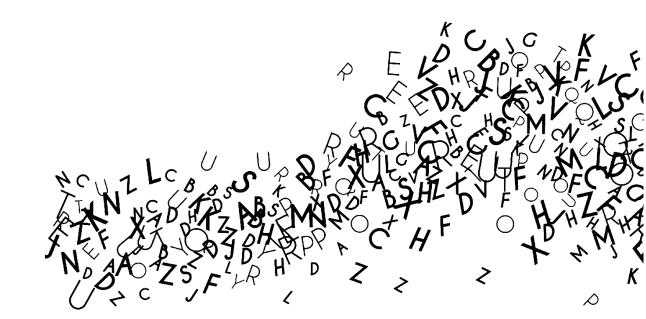


Aim 1: Identifying major life events from adolescence to young adulthood by

- identifying key topics
- emotional valence
- transitions over time

Aim 2: Identifying key themes in young adults' worries and hopes for the future and training a text-based mental wellbeing classifier





Converting text data to numeri- cal format via an embedding model	Language model: 'paraphrase- multilingual-MiniLM-L12-v2'
2. Reducing data <b>dimensionality</b>	UMAP algorithm
'Hard' clustering: Each text entry is categorised into precisely one topic or marked as an outlier	HDBSCAN algorithm
4. Splitting text entries into words ('tokenization')	CountTokenizer (scikit-learn)
Lemmatization, part-of-speech     tagging	spaCy pipeline 'de_dep_news_trf'
Model fine-tuning to avoid topics consisting of highly redundant keywords	Maximal Marginal Relevance criterion
7. Computing <b>importance scores</b> for each word in a cluster	c-TF-IDF algorithm
Merging topics with overlapping content based on previous manual review	BERTopic's 'merge_topics' command
'Soft' clustering assigns probabilities to the entries, allowing membership in overlapping clusters	HDBSCAN algorithm
Updating Topic Representations:     Texts are classified under specific topics or identified as outliers	BERTopic's <i>'update_topics'</i> command
11. <b>Manual check</b> and correction if necessary	Manual Review





**Education &** Career Development

#### School, Education, **Apprenticeship**

beginning apprenticeship education

> studies secondary\_school graduation

matura <sub>school</sub> apprenticeship\_position

#### Work, Employment

job\_position  $j_{\text{professional\_world}}^{\textbf{b}}$ position work military career\_choice entry profession professional\_life Moving & Living on One's Other Life **Transitions &** Independence

## Mental Health &

further\_development music 11 te depression time change

#### **Stressful Life Events** & Loss

accident father illness death grandfather mother grandmother

#### Driver's License & **Vehicles**

Own

shared\_apartment home parental\_home

moving\_out

 $\mathsf{move}_{\mathsf{house}}$ 

residential move-out

driving\_permit car moped driver\_license motorcycle driving\_test driving\_exam driving\_examination

### Wellbeing

#### Mental Health. **Development & Change**

development cliniCpersonality

#### Social Relationships

#### Romantic Relationships & **Friendships**

girlfriend love person friendship boyfriend relationship

#### Family & **Family-Related Events**

mother birth parents sister uncle father brother

#### Marriage & Parenthood

SON pregnancy wedding daughter marriage child engagement mother birth

#### **Leisure Activities** and Successes

#### Travelling, Vacation, Staying Abroad

abroadsummer\_vacation trave⊥ language\_stay exchange\_yearstay\_abroad vacation

family gap\_year month

#### **Sports**

national\_team\_hampionship SUCCESS athleticsjoining soccer basketball

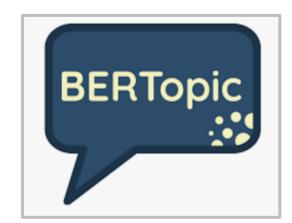
#### **Celebrations & Going Out**

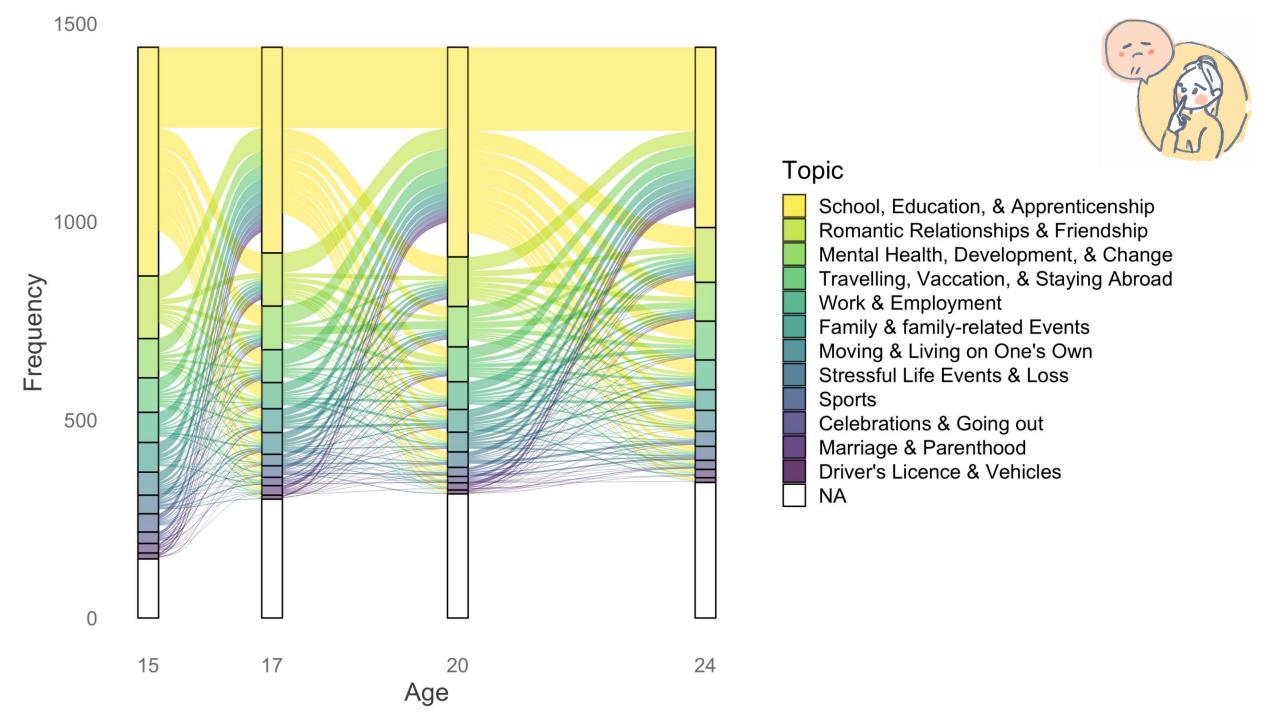
### birthday

Bieber party summer Christmas concert

Justin

5,708 text-entries, 1,442 participants 4 timepoints



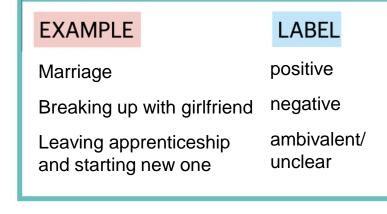


## Classifier training to identify emotional valence



Model interpretability / XAI





New data

Fine-tuned transformer

70% Concern
30% Noconcern

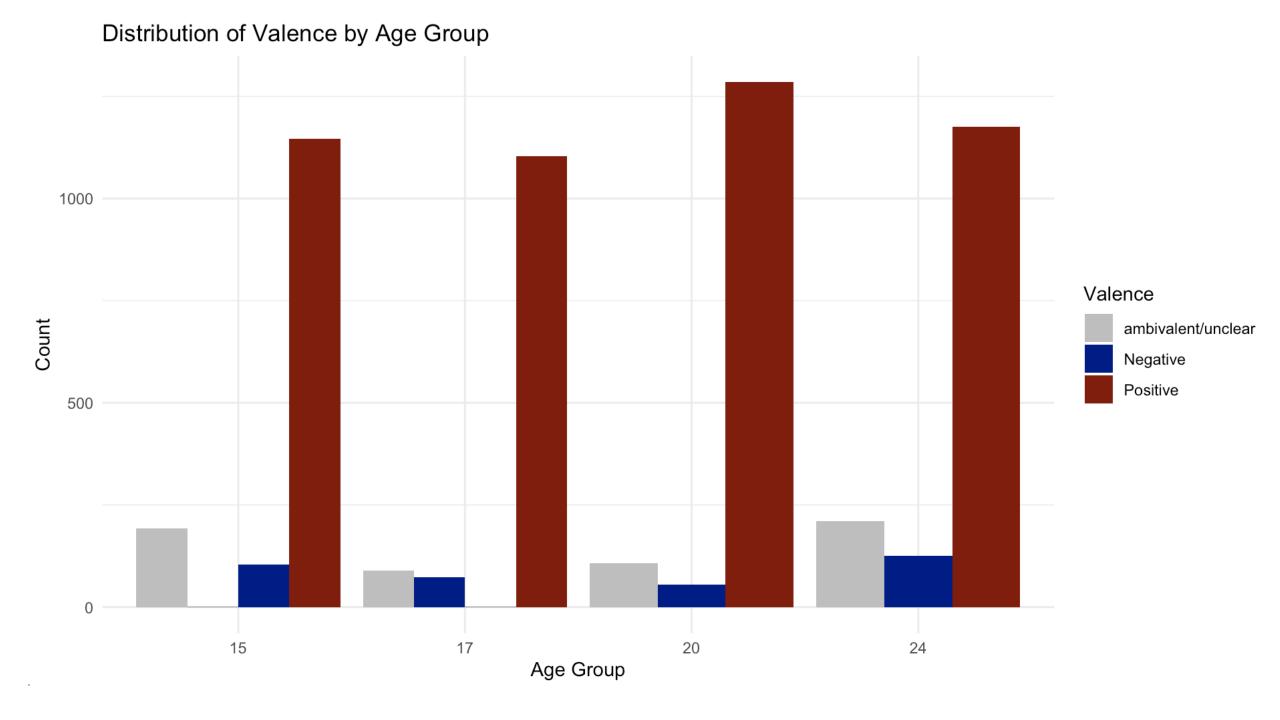
**Pre-trained** 

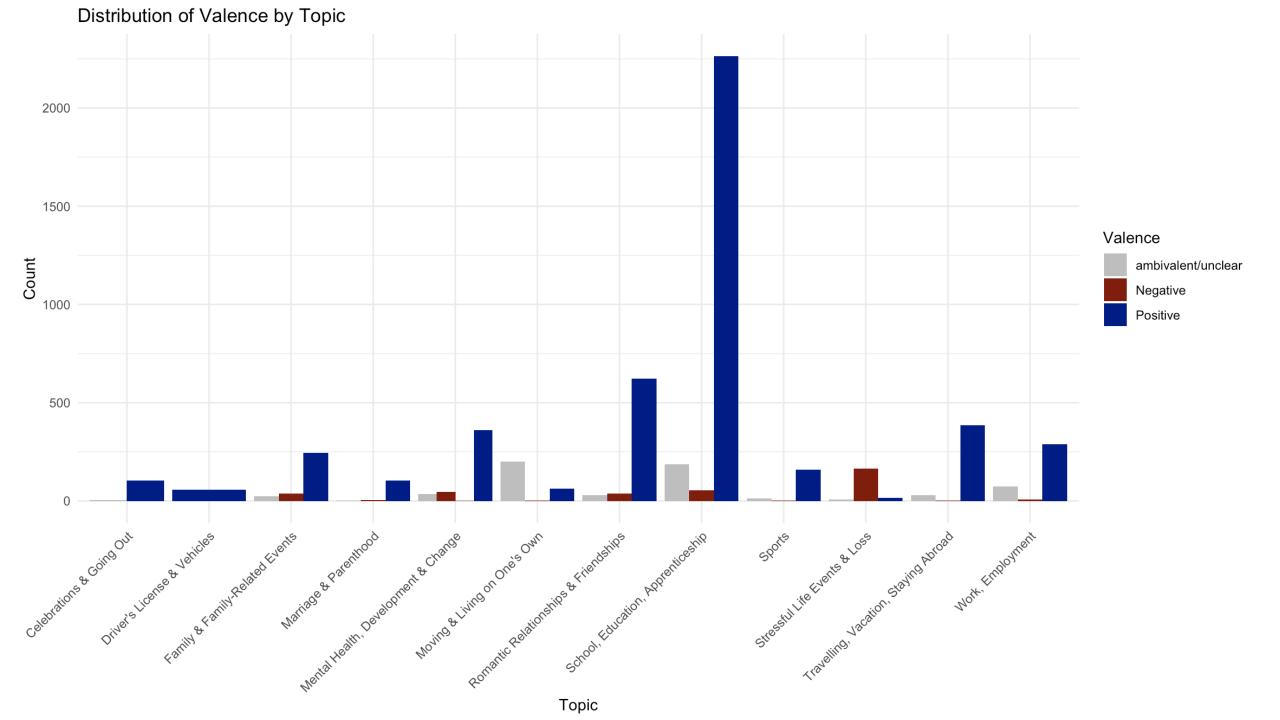
transformer



Image sourced from Demszky et al. (2023), published in Nature Reviews Psychology

**Update parameters** 









```
Word Importance
it was a fantastic performance! #pad
 best film ever #pad #pad #pad #pad
     such a great show! #pad #pad
   it was a horrible movie #pad #pad
i 've never watched something as bad
       that is a terrible movie . #pad
```

## Integrating traditional quantitative methods with automated text analysis can lead to more accurate and personalised insights.

Dr. David Bürgin,
Dr. Christina Haag
&
Lynn Alison Büeler, MSc







Prof. Lilly Shanahan & Prof. Viktor von Wyl

Dr. Denis Ribeaud & Prof. Manuel Eisner

Dr. Laura Bechtiger & Dr. Clarissa Janousch





