

ECALFOR

Bologna meeting





Agenda

- Our team
- NLP Analytics techniques
- NLP applied to ECALFOR data





Learning Analytics Unit

<https://sites.uef.fi/learning-analytics/>



Mohammed Saqr
Associate Professor
Head of LA Unit



Sonsoles López-Pernas
Senior Researcher



Ramy Elmoazen





Research interests

Topics: we work intensively on collaborative learning, engagement, game-based learning, self-regulated learning, computational thinking, problem-based learning, academic achievement, medical education and privacy and ethics.

Our goal: Pushing the boundaries of methodological innovations in learning analytics:

- **Machine Learning** and predictive modeling in general, topic models.
- **Clustering** through dissimilarity-based and model-based techniques.
- **Temporal methods** e.g., Sequence analysis, VaSSTra, Markov models, Multi-channel sequences, Process mining.
- **Network analysis**, e.g., Community detection and Temporal Networks.
- Psychometrics e.g., Psychological networks, Factor analysis, Structural Equation Modeling.





Natural Language Processing

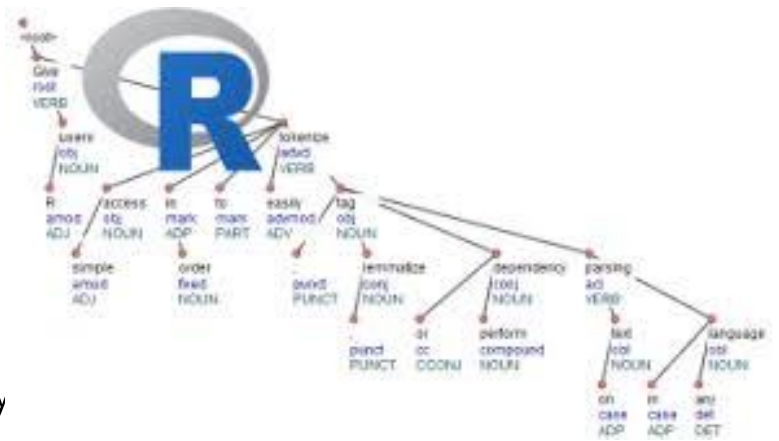
- “The ability of a computer program to understand human language as it's spoken and written”





Text preprocessing

- **Tokenizing:** Dividing the text into smaller units (most commonly words) to facilitate analysis.
- **POS Tagging:** Classifying each token obtained from the previous step according to its grammatical category (e.g., pronoun, verb, adjective, noun).
- **Lemmatizing:** Reducing each word to its dictionary form (e.g., from singular to plural, verb forms to infinitive, etc.).





Sentiment analysis



Fragrance-1
(Lavender)

REVIEWS

1. Smells amazing! A perfect purchase :)
2. Must buy! Super amazing.
3. Quite satisfactory



POSITIVE (81%)



Fragrance-1
(Rose)

REVIEWS

1. A decent purchase
2. Quite okayish! Smells average
3. Could have been better in lot terms



NEUTRAL (88%)



Fragrance-1
(Lemon)

REVIEWS

1. An absolute waste of money.
2. Total waste of money
3. Terrible smell, not worth buytng



negative (91%)

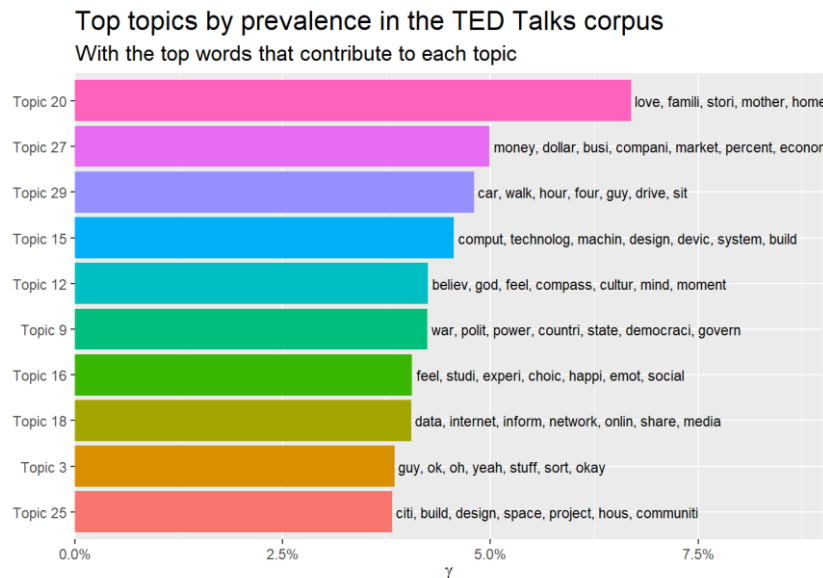
SENTIMENT ANALYZER





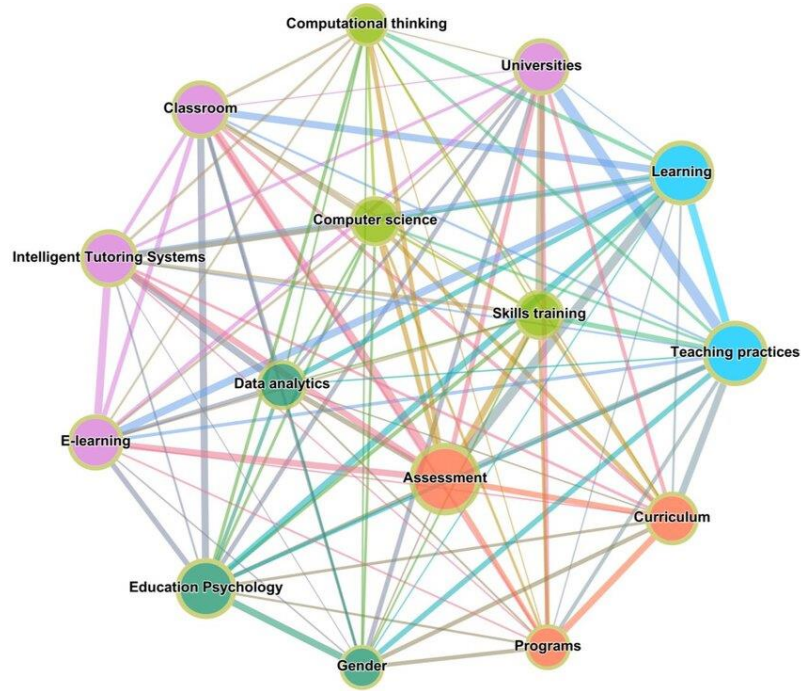
Structured topic modeling

Uncovering latent topics within a corpus of documents.



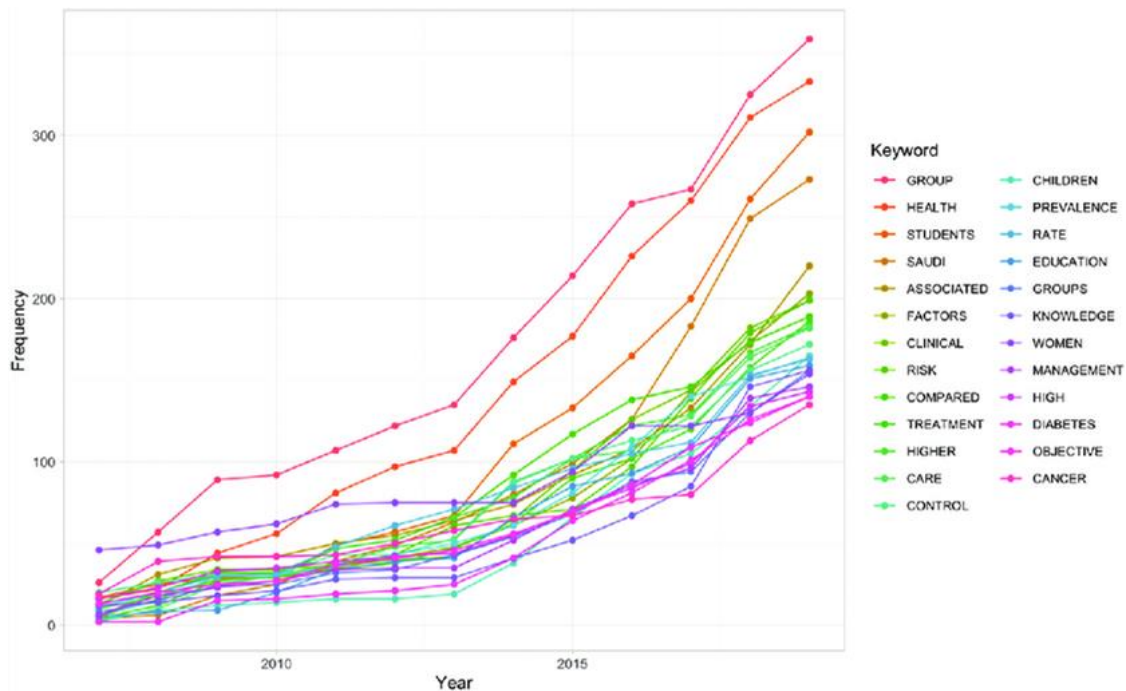


Topic co-occurrence





Topic trends





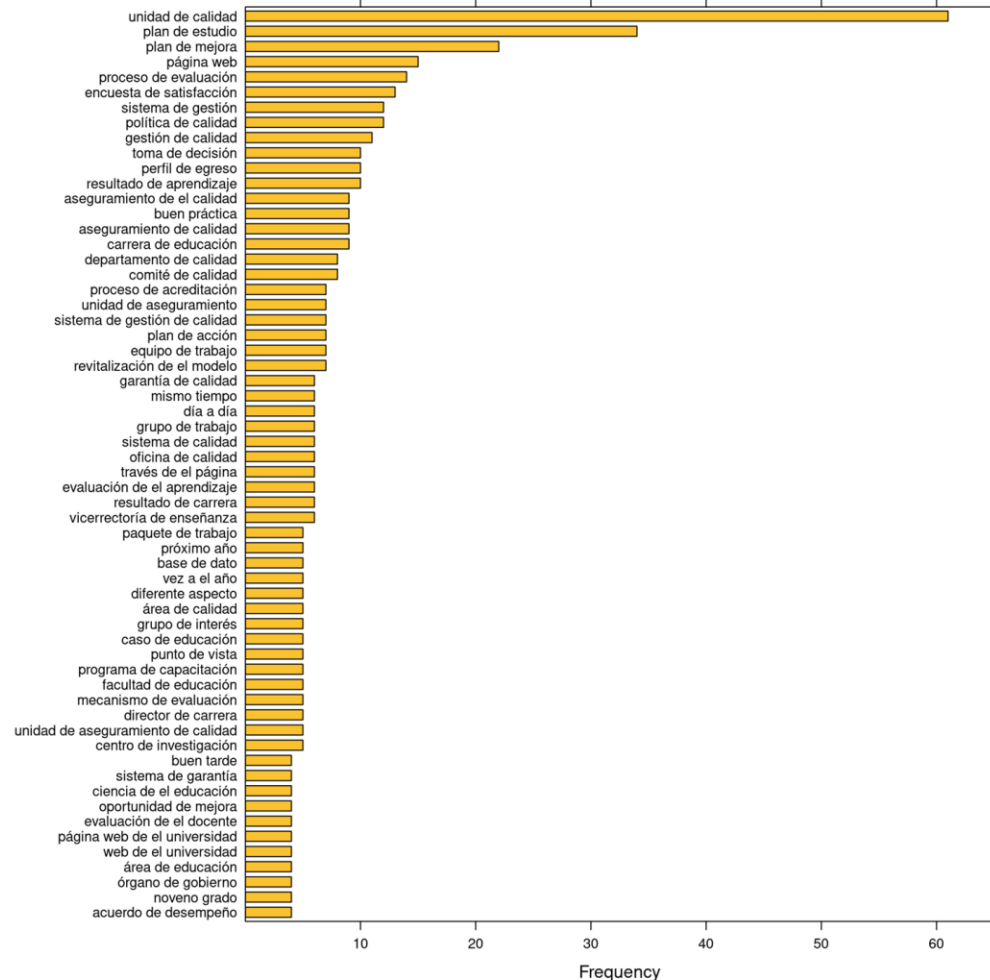
Data

- Interviews from 12 Institutions
- A total of 1983 interactions between interviewer and interviewee
- Transcribed from audio to Word file
- Prepared into Excel format to enable processing





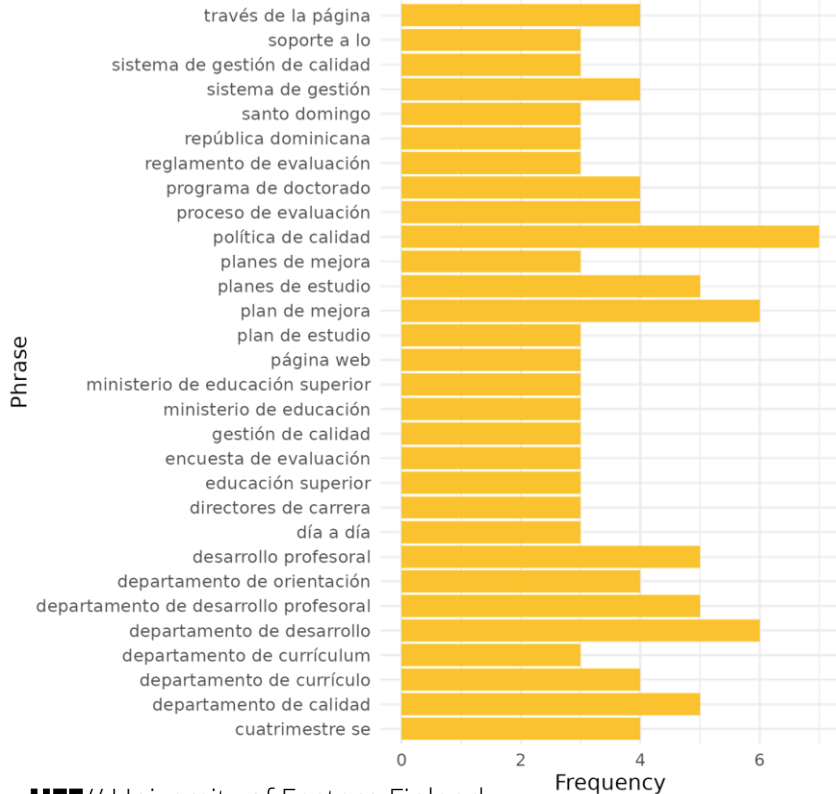
Phrases



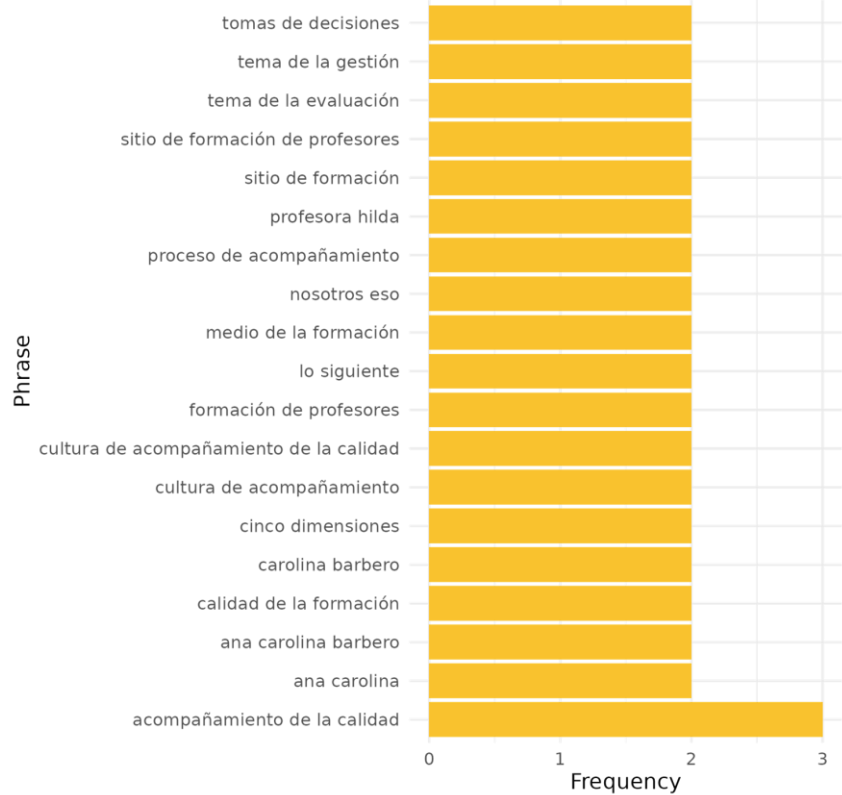


Phrases

UTESA



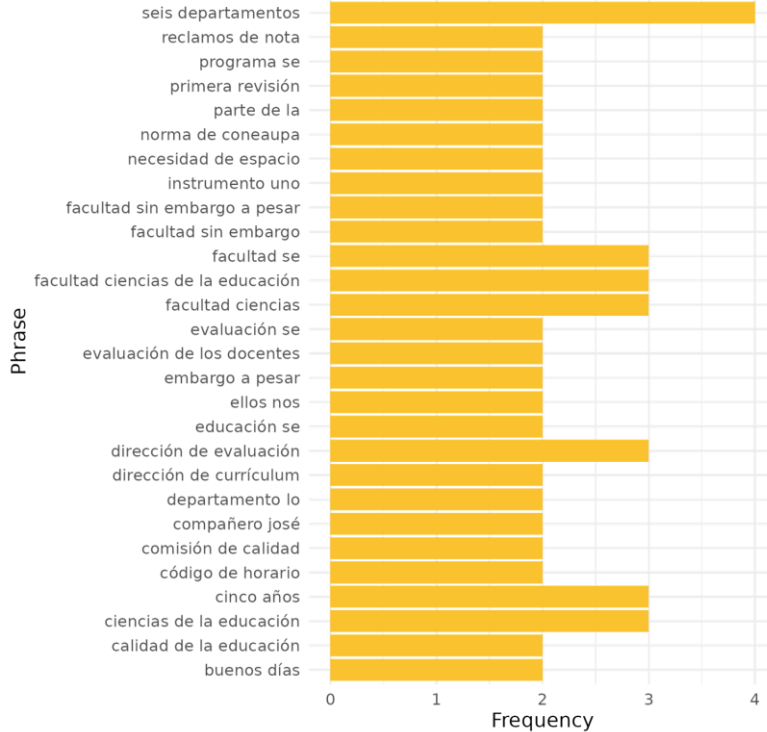
UNESP



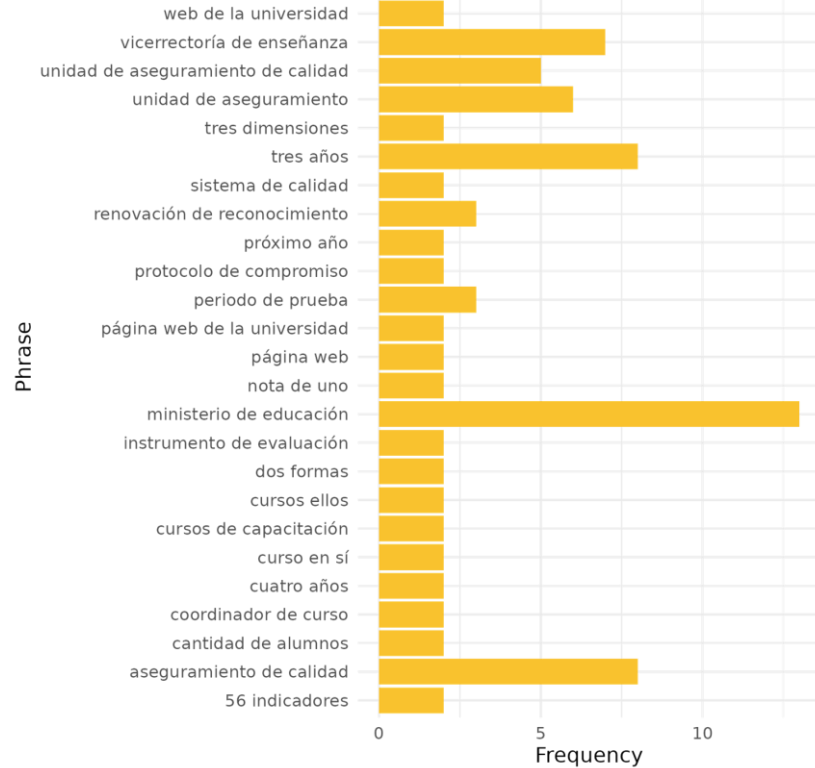


Phrases

UNACHI

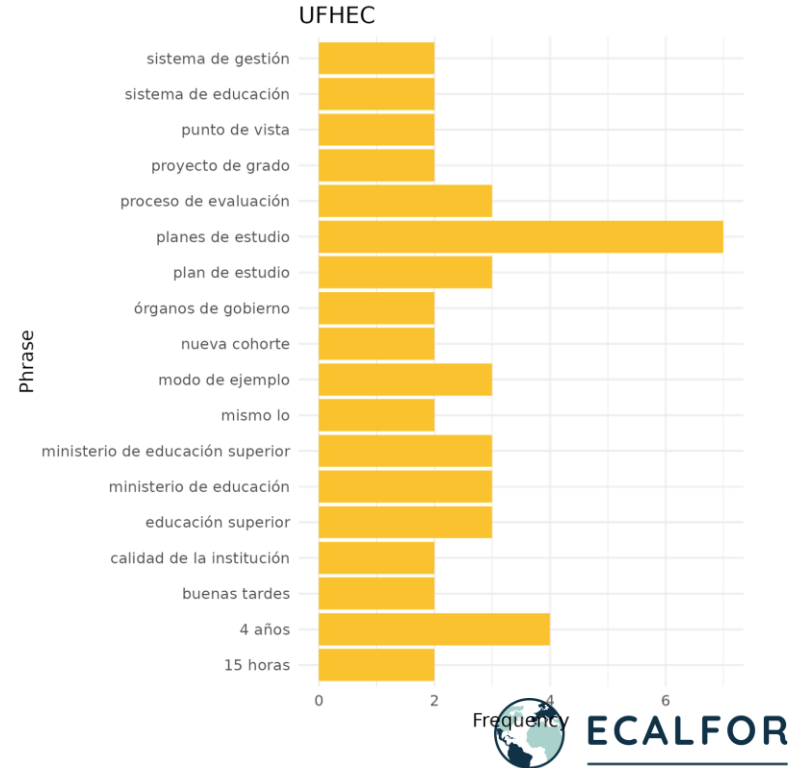
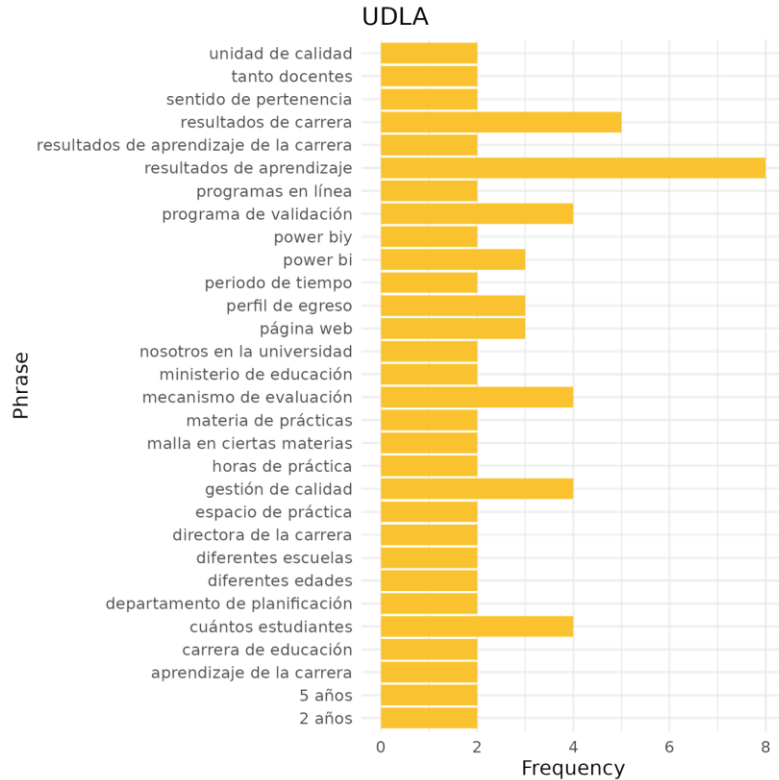


UFMA



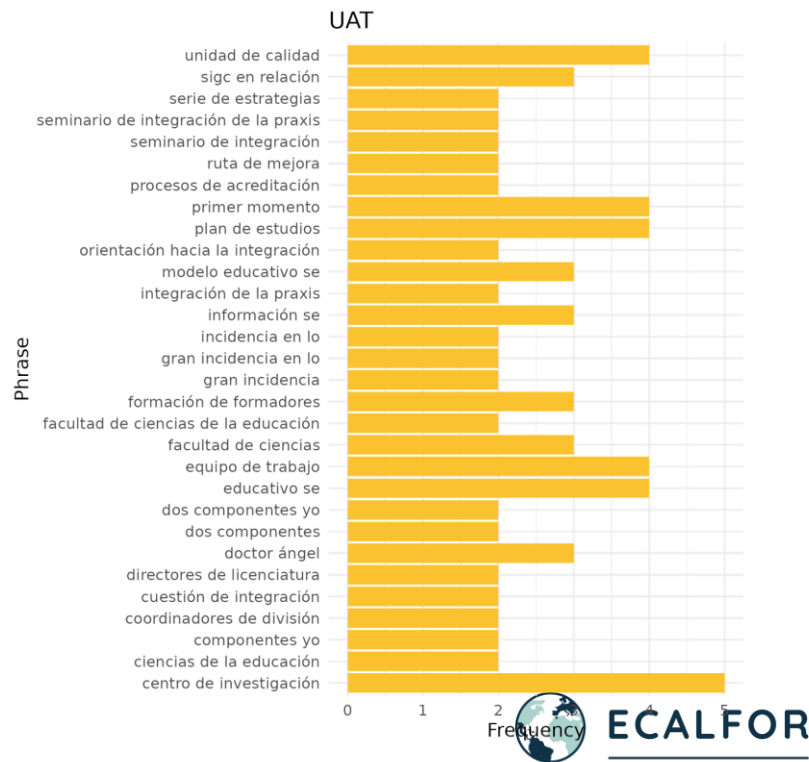
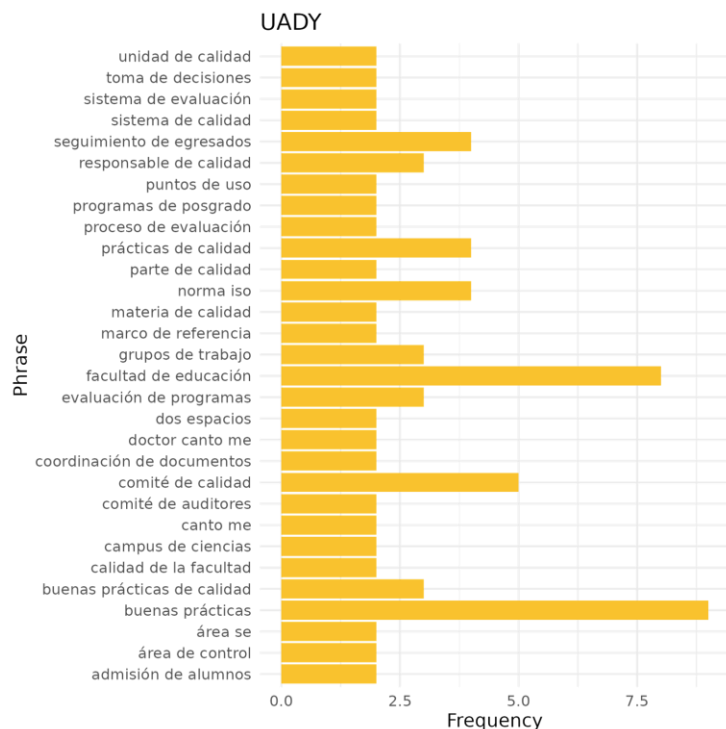


Phrases



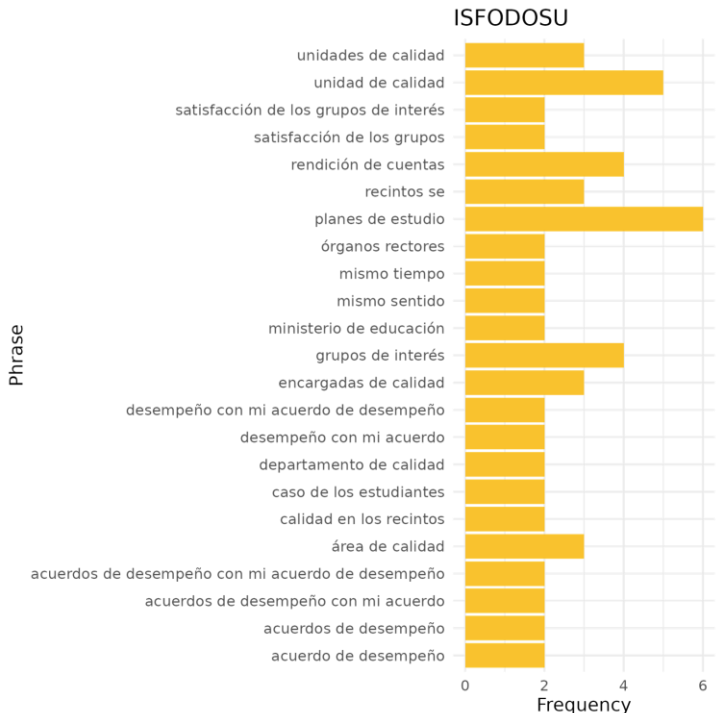
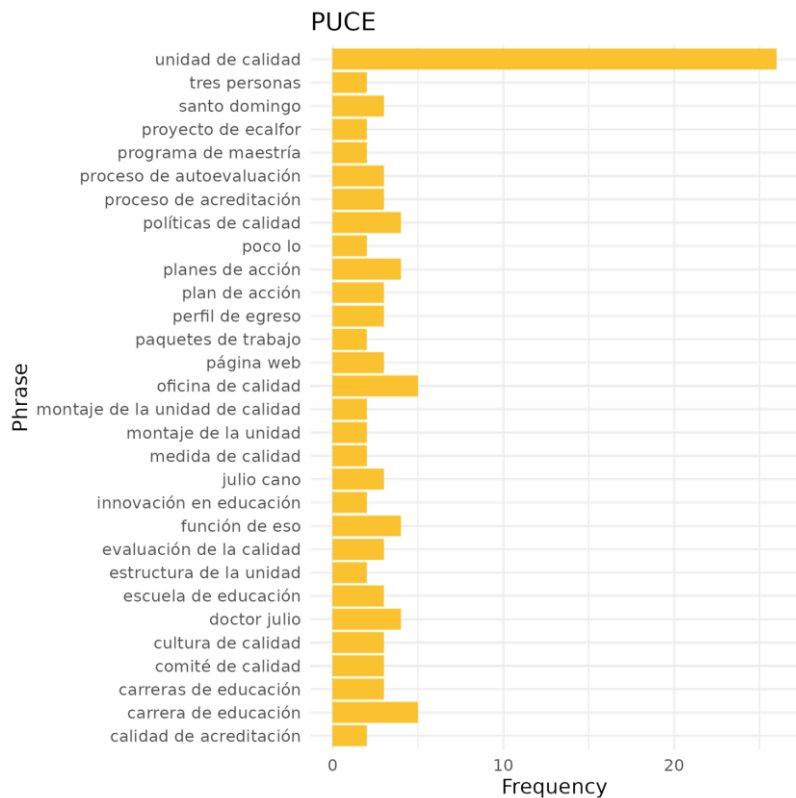


Phrases



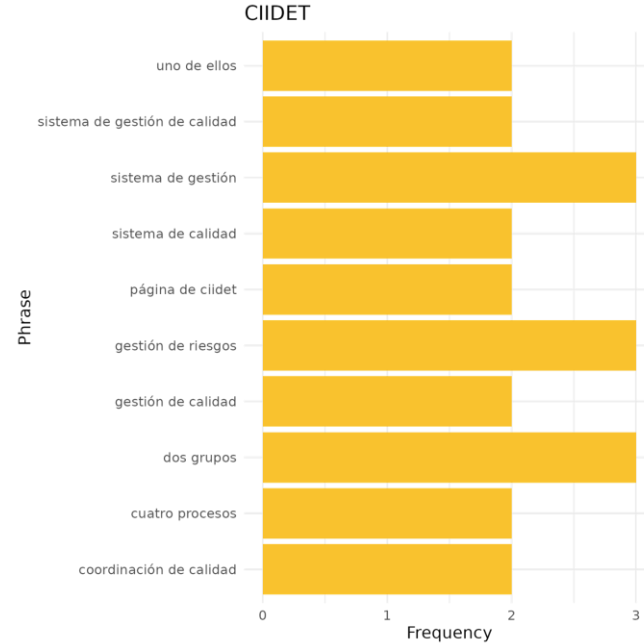
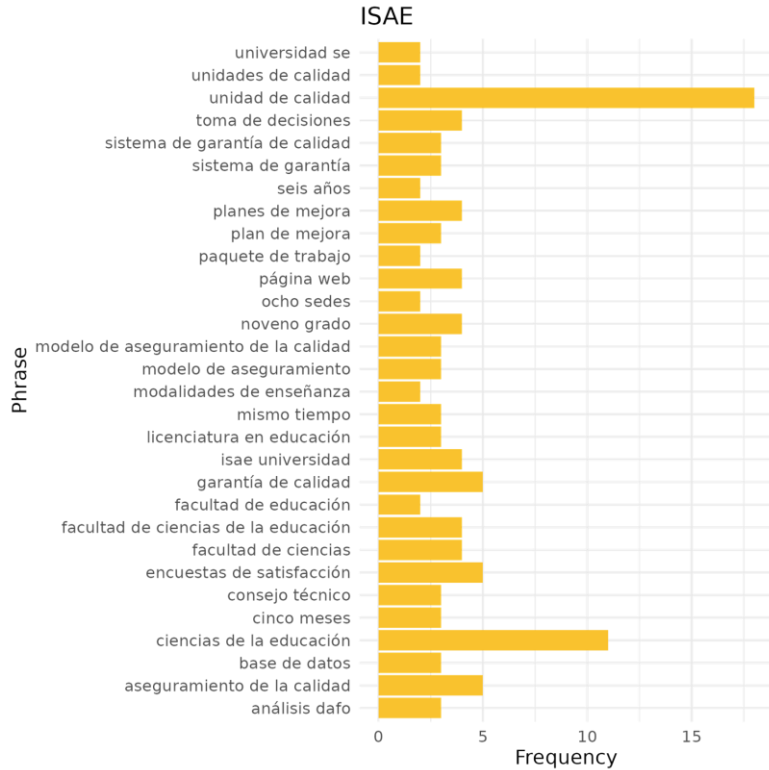


Phrases

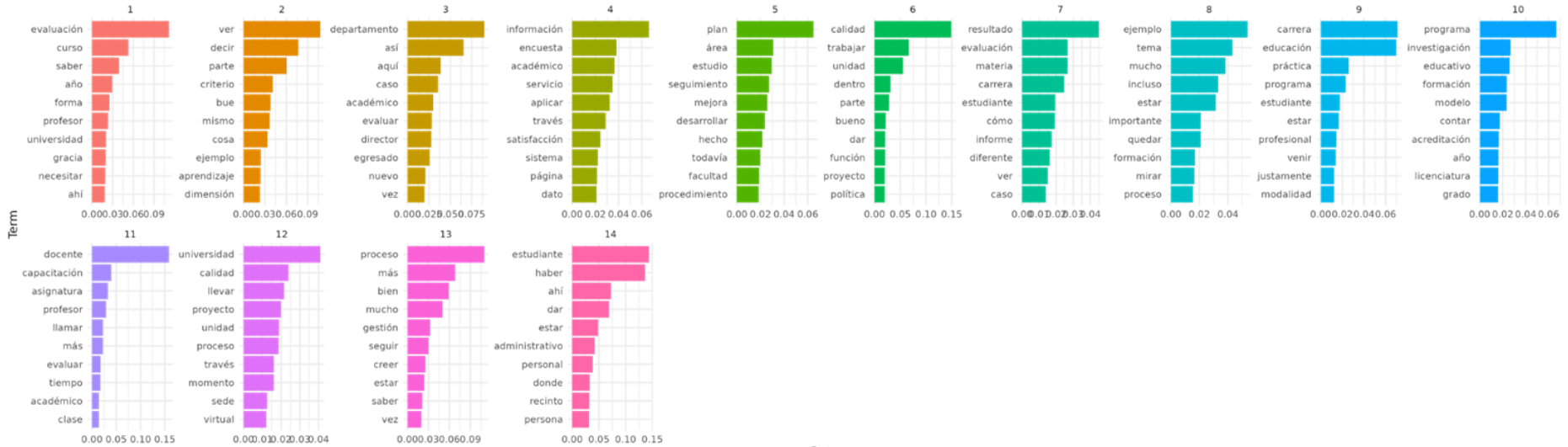




Phrases



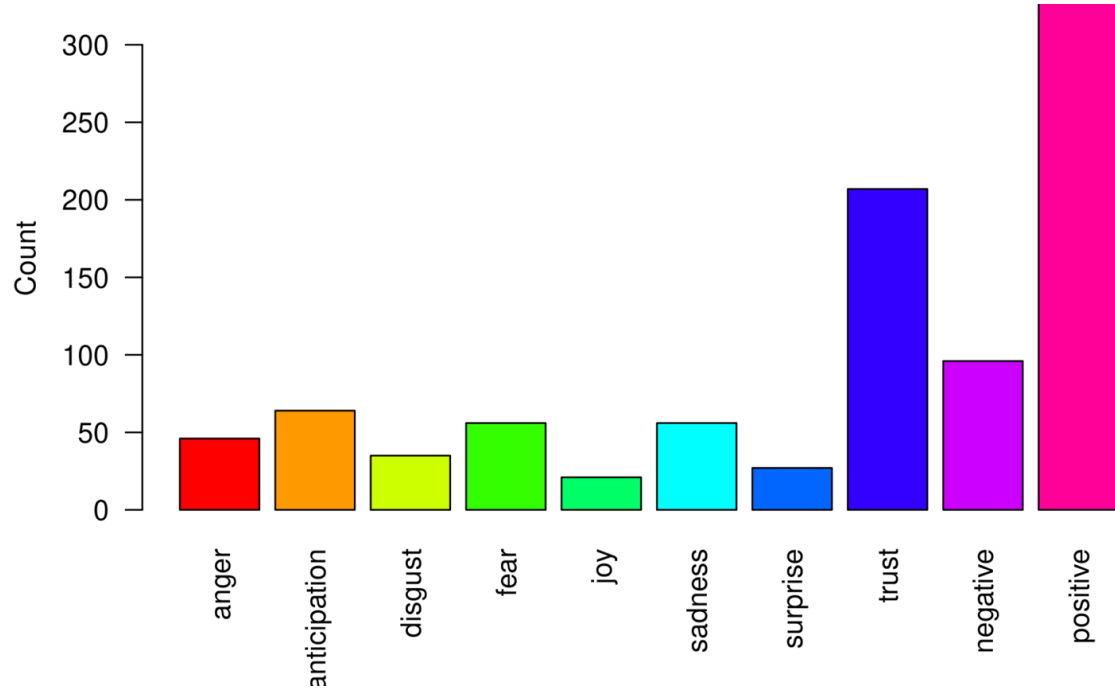
Topics



Beta



Sentiment analysis



Sentiment analysis



Examples

- <https://ieeexplore.ieee.org/document/9729804>
- <https://www.sciencedirect.com/science/article/abs/pii/S1041608023001437>
- <https://www.sciencedirect.com/science/article/pii/S0747563221004465>
- <https://www.sciencedirect.com/science/article/pii/S0747563223004272>



Action points

- Code data qualitatively (NVivo)
- Networks
- Correlation

