



Concept

 Producing Alignment Data by "Citizen Scientist"

 Win-Win Situation – Learning Effect on Users





What Texts are Used

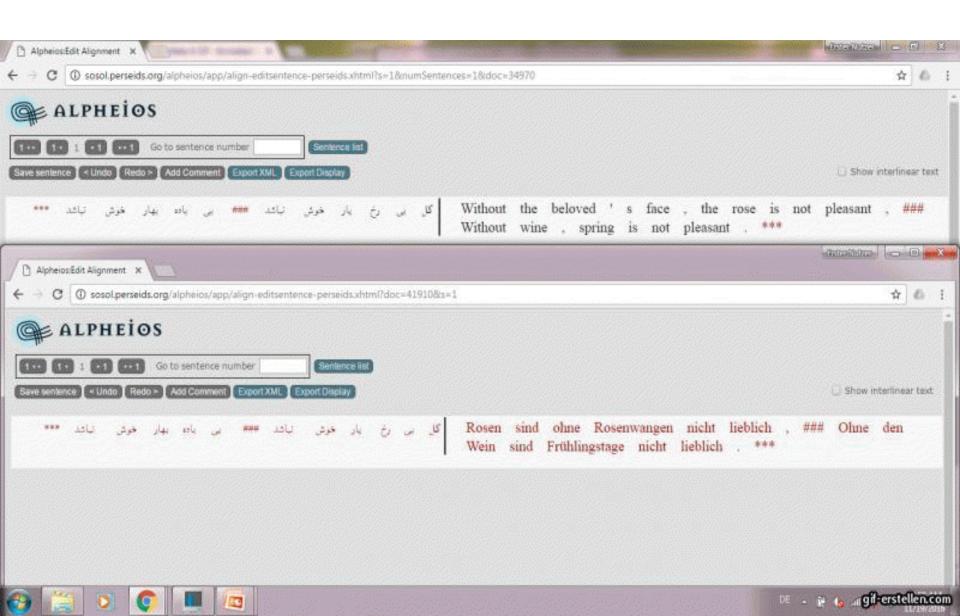
Divan-e-Hafez as the source Text

The English translation of Colonel
 Wilberforce Clarke, published in 1891

 The German translation of Joseph von Hammer-Purgstall, published in 1812



How It Works





How It Doesn't Work

فریاد که از شش جهتم راه ببستند

آن خال و خط و زلف و قد و عارض و قامت

Justice For, they have barred my Path on <u>six</u> sides
By the power of that mole, beard, tress, face, cheek,
and stature

Wehe! wehe! der Weg ist verrennt von allen <u>vier</u> Seiten,

Durch die Locken, durchs Maal, durch das Gesicht und den Wuchs.



And Still, How It Doesn't Work



ترک عاشق کش من مست برون رفت امروز تا دگر خون که از دیده روان خواهد بود

Today, my Bold one, lover-slayer, went for thin toxicated,

Let us see again from whose eye, blood-flowing shall be

Es gieng mein Vielgeliebter, Heut auf den Raub aus, Wem wohl's Loos zu bluten, Bestimmet seyn wird.



What We Can See with Alignment



• Cultural Features (<u>تاریک</u>)

 Comparing translations and not reading just ONE translation (پسر)

Linguistic Features (<u>J</u>)

Asking New Questions (شراب)





User's Observation

- Genderlessness of 3rd Person Singular Pronoun
- No Articles in Persian
- Position of the Predicate
- Shorter Form of Some Words
- Position of the Adjective and Possessive Pronoun





HINT!

DON'T use translation alignment if you have a vocabulary quiz tomorrow!



Vocabulary Learning Assessment



Comparing with digital flash cards

Vocabulary Knowledge Scale

In Two Steps: With and Without Context

 Running Delayed Test after 2 Weeks and 2 Months





Future Work

 Adding Morphological and Syntactical Annotation

Adding Audio





Thank you!

@OpenPersian maryam.foradi@uni-Leipzig.de

