



# Development of an Asymptotic Word Correspondence System between Classical Japanese Poems and their Modern Translations

Hilofumi Yamamoto\*<sup>†</sup>  
\* University of California, San Diego

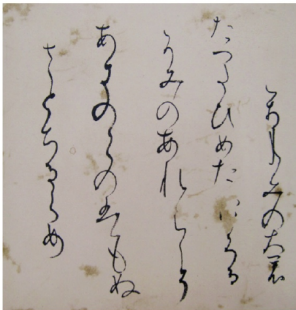
Hajime Murai<sup>†</sup>  
<sup>†</sup>Tokyo Institute of Technology

Bor Hodošček<sup>‡</sup>  
<sup>‡</sup>Meiji University

## Objectives:

The development of the **THESAURUS** of Classical Japanese Poetry. (10–13th century; Heian period )

### Waka: Japanese poetry



*Tatsuta-Hime...*  
*tamukuru kami no / arebakoso*  
*aki no konoha no / nusa to chirurame*  
because princess Tatsuta has a god whom she offers  
brocades, the leaves of autumn wil scatter as an offering.



## Problems:

- Unit: 卵の花 or 卵 / の / 花
- Orthography: sad!  
さびしい / さみしい / 寂しい / 淋しい
- Attribution:  
卵の花 ∈ plant or 卵の花 ∈ food



## Solution:

- Using Pararell Corpora:  
Original Poems and Contemporary Translation Texts
- Using Mutual Co-occurrence Ratio:  $mcr(o, t) = p(o|t)p(t|o)$   
To estimate the rate of contextual similarity between classical words and comtemporary words.