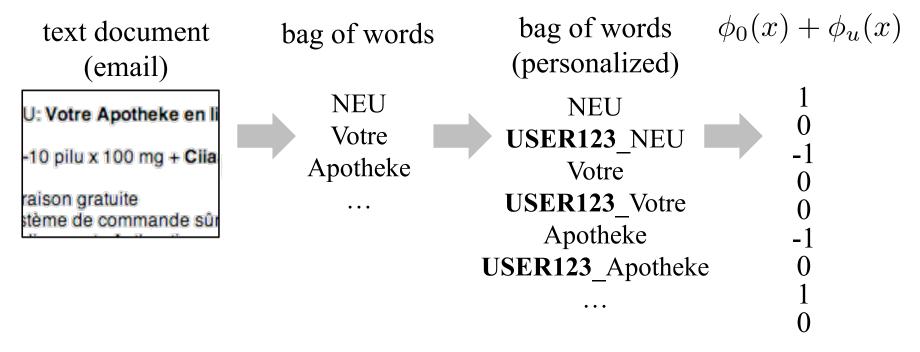
## Trillion features with hashing

## Personalized tokens trick

- $\phi_o(token) = hash(token) \% 2^b$
- $\phi_u(token) = hash(u + "\_" + token) \% 2^b$
- We obtain 16 trillion pairs (user, word) but still  $2^b$  features



https://arxiv.org/pdf/0902.2206.pdf