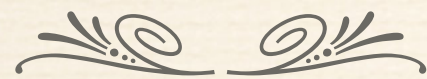


Speak 'Friend' and Enter: Textual Analysis of The Lord of the Rings



Data Science by John D Hazard

'The Lord of the Rings' by J.R.R. Tolkien

Artwork by Ted Nasmith



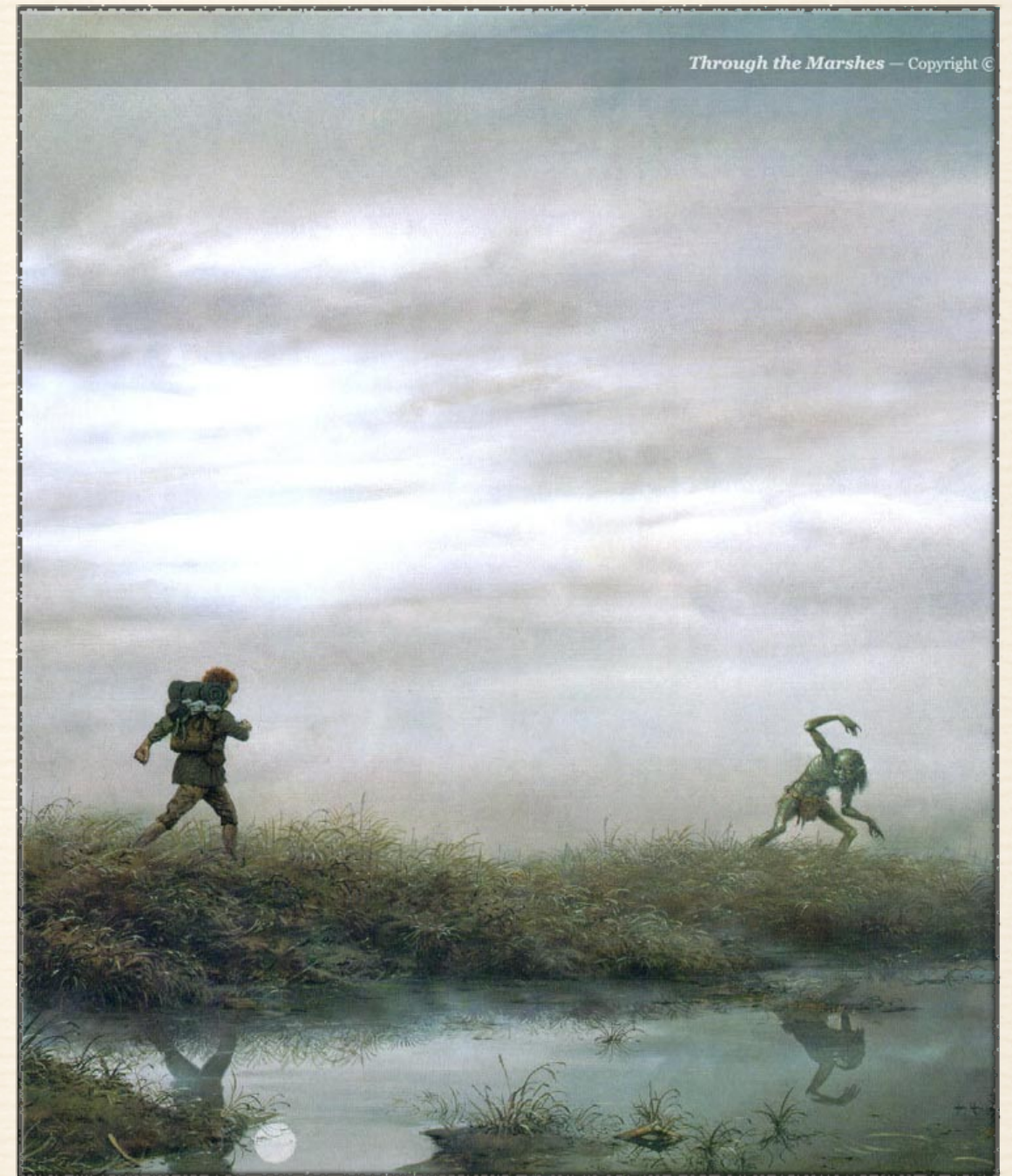
*“Why should a man be scorned if, finding himself in prison, he tries to get out and go home? Or if, when he cannot do so, he thinks and talks about other topics than jailers and prison-walls? The world outside has not become less real because the prisoner cannot see it. In using escape in this way the critics have chosen the wrong word, and, what is more, they are confusing, not always by sincere error, the *Escape of the Prisoner* with the *Flight of the Deserter*. ”*

—J.R.R. Tolkien from ‘On Fairy Stories’

What is Topic Modeling?

Topics From Latent Dirichlet Allocation

- ❖ Top Words: 'great, end, turned, wind, path, friend, ride, drew, tale, mile, plain, precious, laughed, life, sprang, remember, swiftly, climbed, dim, flame'
- ❖ Top Words: 'away, black, passed, gollum, fell, slowly, field, power, hold, sent, bit, grass, tirith, feel, hurt, mouth, aside, falling, desire, finger'



Topics From Non-Negative Matrix Factorization

- ❖ Top Words: ‘pippin, shadow, think, good, king, made, suddenly, two, head, deep’
- ❖ Top Words: ‘saw, ring, behind, good, stood, shadow, end, black, ever, foot’
- ❖ Top Words: ‘gandalf, end, ever, black, foot, stone, passed, left, hill, much’



What is Word2Vec?

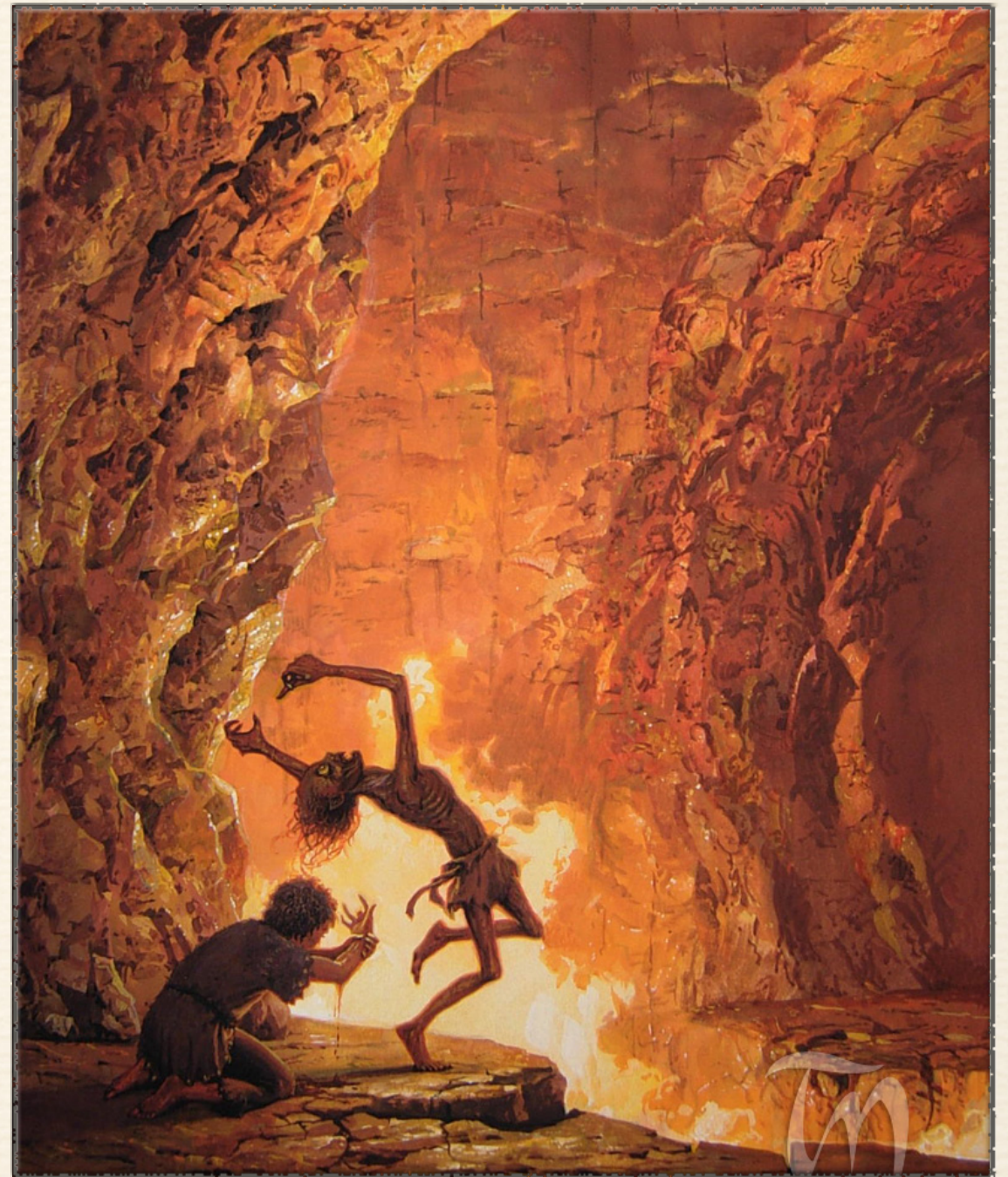
Character Study: Tom Bombadil

- ❖ CBOW Most Similar:
maggot 0.97, shouting 0.97,
hints 0.97
- ❖ SKIP-GRAM Most Similar:
maggot 0.92, party 0.93,
gandalf 0.93



Object Study: The Ring

- ❖ CBOW Most Similar: enemy 0.89, journey 0.88, company 0.87
- ❖ SKIP-GRAM Most Similar: enemy 0.83, burden 0.80, council 0.79



Place Study: The Shire

- ❖ CBOW Most Similar:
ringwraiths 0.90, enemy
0.88, war 0.88
- ❖ SKIP-GRAM Most Similar:
nine 0.87, truth 0.87, middle-
earth 0.87





There And Back Again