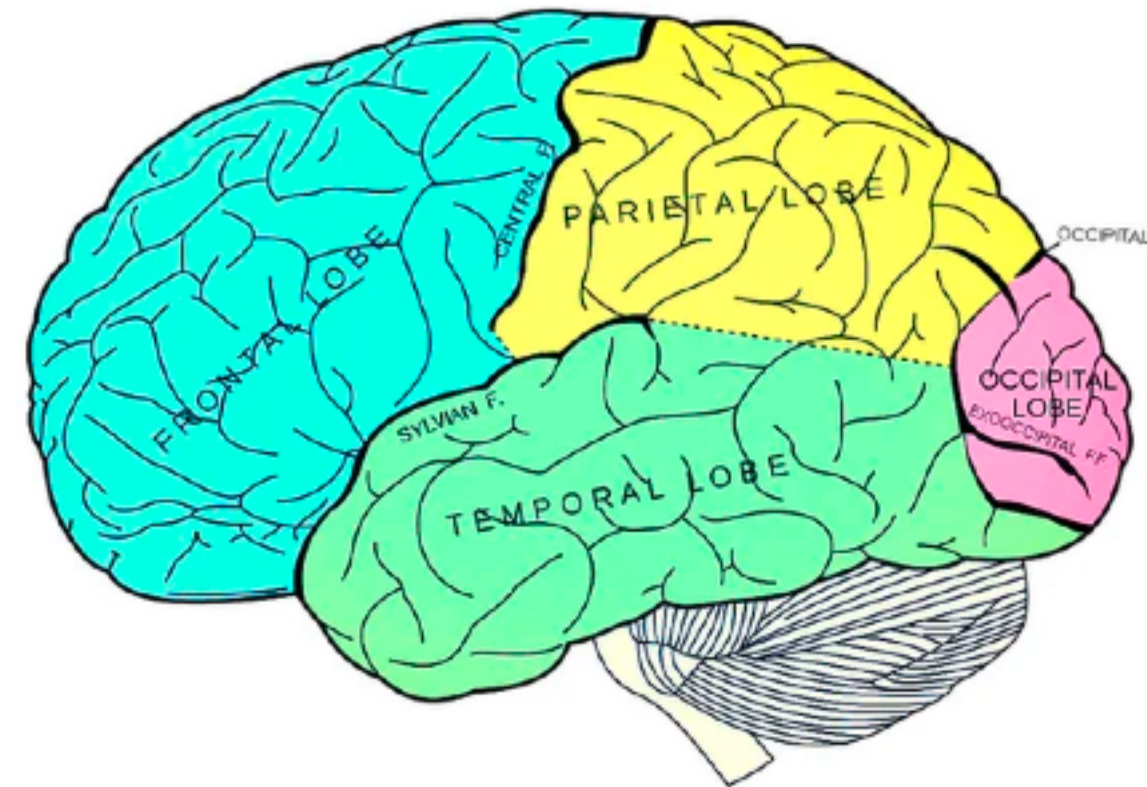# What is Natural Language Processing (NLP)

- Closely related to the fields of computational linguistics and machine learning

  - Linguistics is concerned with language itself; its structure and how we represent it, the nature of meaning derived from it, and the application of logic and reasoning over it.

  - In contrast, the machine learning approach primarily aims to build computational methods that enable computers to processes human language with the goal of performing some task. Examples include search, question answering, conversing with humans etc.

- Modern NLP relies most heavily on the machine learning approach.

The machine learning approach

At a high level, machine learning is set of techniques, or *machinery*, that is designed to automatically learn complex relationships between random variables. Typically there is some output that we are interested in predicting, and there are inputs that we hypothesize affect that output. The goal is accurately model real world phenomena.
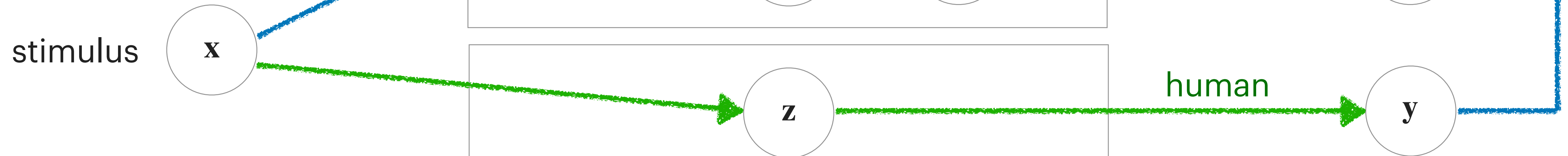
In the case of NLP, the underlying data generating process lies in the realm of human cognition, which practically speaking is a black box.

surrogate

$f(\mathbf{x};\boldsymbol{\theta})$

$\mathbf{L}(\hat{\mathbf{y}}, \mathbf{y}, \boldsymbol{\theta})$

$\hat{\mathbf{z}}$

$g(\hat{\mathbf{z}})$

machine

$\hat{\mathbf{y}}$

stimulus

$\mathbf{x}$

$\mathbf{z}$

human

$\mathbf{y}$

unobserved hidden state