

# Praktikum Kecerdasaan Buatan

## Natural Language Processing (NLP)

Renovita Edelani

Entin Martiana Kusumaningtyas

Yuliana Setiowati

Ali Ridho Barakbah

2024



Politeknik Elektronika Negeri Surabaya  
Departemen Teknik Informatika dan Komputer

# Overview

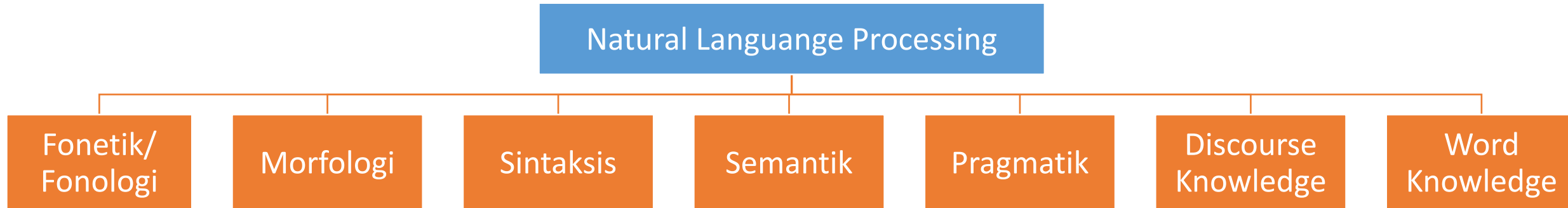
- Dasar Teori
- Percobaan Praktikum
- Tugas Praktikum

# Natural Language Processing

Proses **pembuatan model komputasi** dari **Bahasa** sehingga memungkinkan terjadinya **interaksi antara manusia dan computer** dengan **perantaraan Bahasa alami** yang dipakai oleh manusia

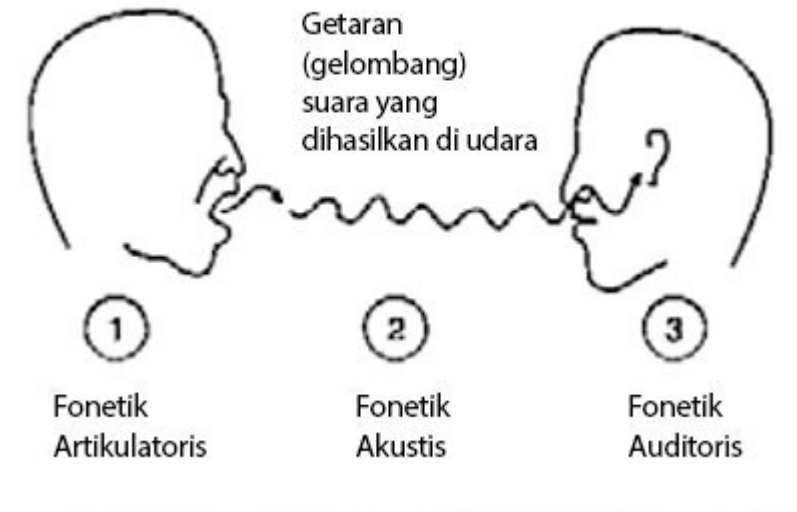
NLP **memodelkan pengetahuan** terhadap **Bahasa**, baik dari segi kata, bagaimana kata-kata bergabung menjadi suatu kalimat dan konteks kata dalam kalimat

# Disiplin Ilmu dari NLP



# Fonetik/Fonologi

- Berhubungan dengan **suara** yang menghasilkan kata yang dapat dikenali
- Dipakai dalam **Speech based System**



# Morfologi

- Pengetahuan tentang **kata dan bentuknya** sehingga **bisa dibedakan** antara yang satu dengan yang lain.
- Contoh:
  - membangunkan
    - bangun (kata dasar)
    - mem (prefix)
    - kan (suffix)

# Sintaksis

- Pengetahuan tentang **urutan kata** dalam **pembentukan kalimat**.

- Contoh:

Kalimat → **Subyek**, **Predikat**

Kalimat → Subyek, Predikat, Keterangan

**Subyek** → **Determinan**, **KataBenda**

**Predikat** → **KataKerja**, **KataBenda**

**Determinan** → seorang, seekor

**KataBenda** → manusia, anjing, nasi, daging, air

**KataKerja** → makan, minum

**Keterangan** → kemarin, tadi

Subyek □ seorang, manusia

Predikat □ makan, nasi

Kalimat □ seorang, manusia, makan, nasi

# Semantik

- Mempelajari **arti suatu kata** dan bagaimana arti kata-arti kata tersebut membentuk suatu arti kata dari kalimat yang utuh.
- Contoh:
  - Ayahku datang membawa **buah tangan**
  - Budi terkenal **panjang tangan**



# Pragmatik

- Pengetahuan tentang **konteks** kata/kalimat yang berhubungan erat keadaan atau situasi kata/kalimat tersebut dipakai.
- Contoh :
  - Ayah datang (diucapkan dengan **nada datar**)
  - Ayah datang! (diucapkan dengan **nada tinggi**)
  - Ayah datang? (diucapkan dengan **tempo cepat**)

# Discourse Knowledge

- Pengetahuan tentang **hubungan antar kalimat**.
- Melakukan pengenalan apakah suatu **kalimat yang telah dikenali mempengaruhi kalimat selanjutnya**.
- Penting untuk **identifikasi kata ganti orang, keterangan tempat atau aspek sementara dari informasi**.

- Contoh:

Ibu pergi ke pasar. Ia membeli makanan disana.



# Word Knowledge

Mencakup **arti sebuah kata secara umum** dan **apakah arti khusus** bagi suatu kata dalam suatu percakapan dengan **konteks tertentu**

Ia **menangkap** tikus itu

A: Apakah kamu memahami kuliah tadi?  
B: Saya sudah **menangkap**

# Aplikasi NLP

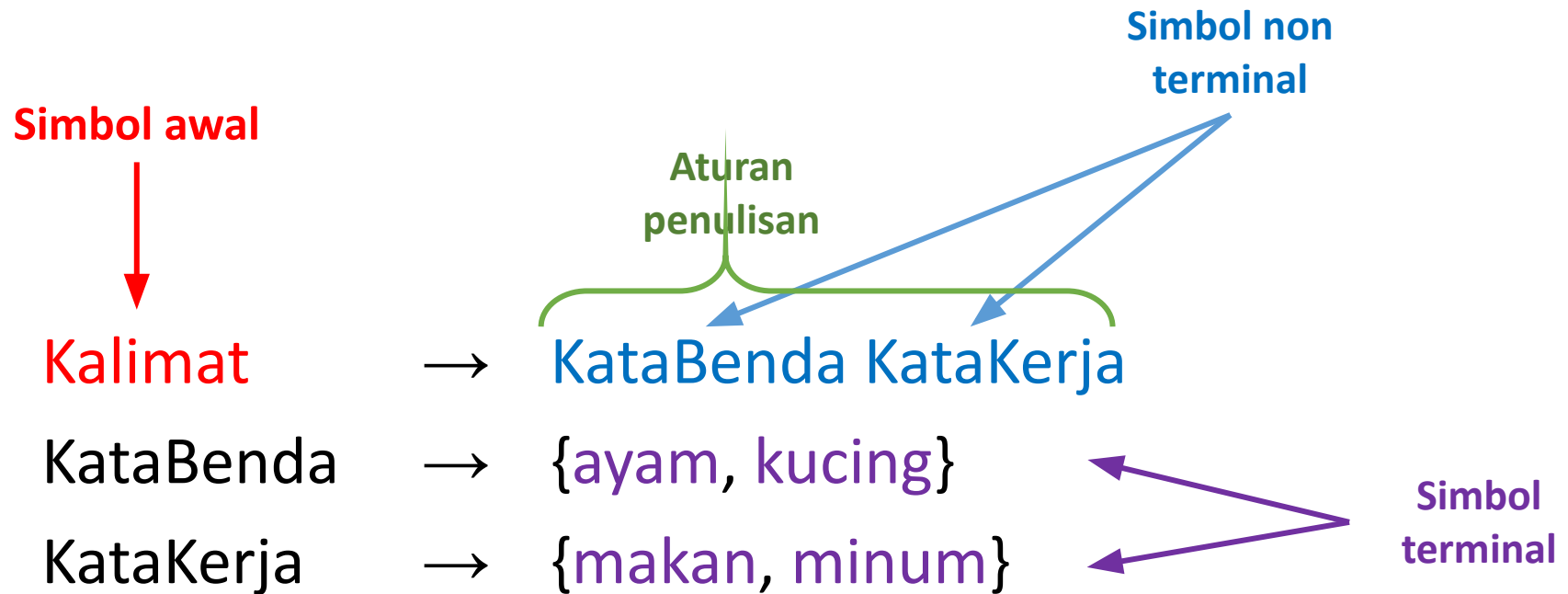
Text-based Application	Speech-based Application
<p data-bbox="104 444 1156 551">Aplikasi yang melakukan memprosesan terhadap <b>teks tertulis</b></p> <p data-bbox="104 634 275 675">Contoh:</p> <ul data-bbox="206 696 1225 1053" style="list-style-type: none"><li data-bbox="206 696 1021 803">• Mencari topik tertentu dari buku di perpustakaan</li><li data-bbox="206 822 1116 863">• Mencari isi dari suatu berita atau artikel</li><li data-bbox="206 882 721 923">• Mencari isi dari email</li><li data-bbox="206 942 1225 1053">• Menerjemahkan dokumen dari suatu Bahasa ke Bahasa lain</li></ul>	<p data-bbox="1268 444 2372 551">Aplikasi yang melakukan memprosesan dari <b>Bahasa lisan atau pengenalan suara</b></p> <p data-bbox="1268 634 1439 675">Contoh:</p> <ul data-bbox="1370 696 2390 868" style="list-style-type: none"><li data-bbox="1370 696 2359 738">• Sistem otomatis pelayanan melalui telepon</li><li data-bbox="1370 756 2283 798">• Kontrol suara pada peralatan elektronik</li><li data-bbox="1370 816 2390 868">• Aplikasi peningkatan kemampuan berbahasa</li></ul>

# Gramatika

- Suatu **aturan** yang menentukan apakah **suatu kumpulan kata dapat diterima sebagai kalimat** dalam suatu Bahasa
- **Context Free Grammar (CFG)** adalah representasi grammar dari Chomsky Hierarchy yang mudah dipahami dan diolah dalam bentuk program.
- CFG merepresentasikan bahwa suatu grammar itu dapat dibentuk dari 4 elemen, yaitu:
  - Simbol awal
  - Aturan penulisan
  - Simbol non terminal
  - Simbol terminal



# Gramatika



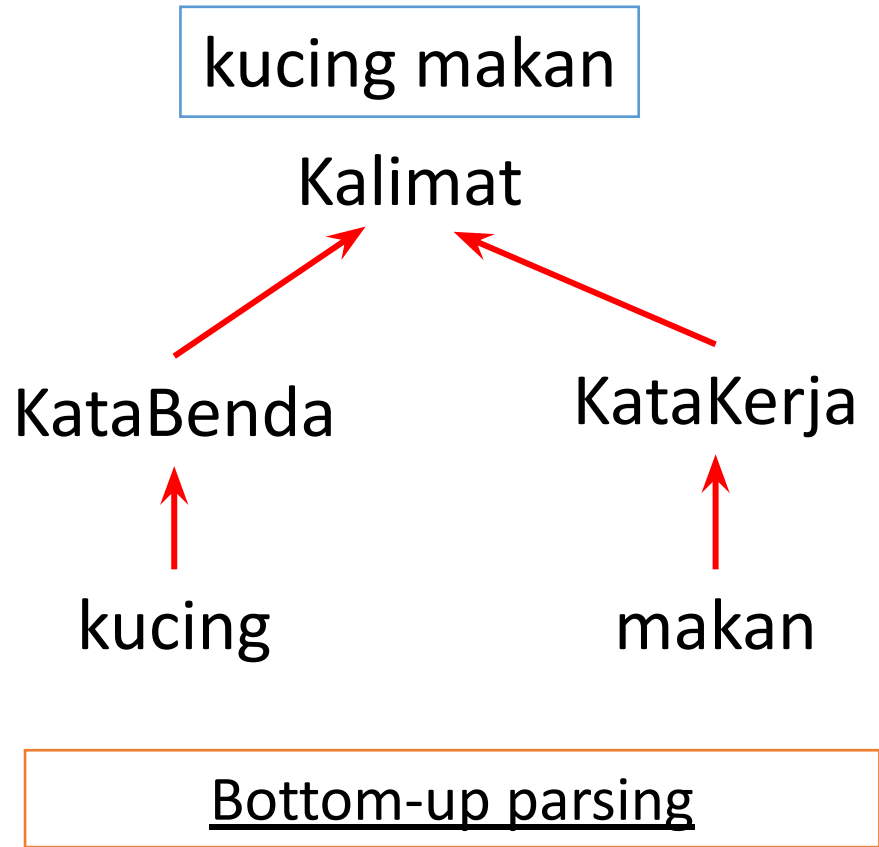
# Parsing

Suatu **proses menganalisa suatu kumpulan kata** dengan **memisahkan kata-kata** itu dan **menentukan struktur sintaksis** dari tiap kata tersebut

Mempunyai 2 pendekatan:

- Top-down parsing
- Bottom-up parsing

# Parsing





# Hands on Lab Practice

Access via Link: <https://pens.id/NLP2024>



# Tugas Praktikum 1

1. Ambil satu berita online Bahasa Indonesia yang nantinya akan dijadikan data. (**Setiap mahasiswa harus berbeda** beritanya)

$text \leftarrow$  satu berita online Bahasa Indonesia

2. Lakukan proses tokenizing *word\_tokenize* untuk *text*

$tokenized\_word \leftarrow$  text yang sudah di tokenize perkata

3. Remove/Buang kata stopwords Bahasa Indonesia yang ada pada *tokenized\_word*

$i\_filtered\_list \leftarrow$  array of kata yang sudah di filter stopwords-nya

4. Lakukan proses Stemming pada *i\_filtered\_list* menggunakan library Sastrawi

$i\_stemmed\_words \leftarrow$  kata yang sudah di-stemming menggunakan Sastrawi

5. Hitung distribusi frekuensi dari kata pada *i\_stemmed\_words*



## Tugas Praktikum 2

- Rubahlah CFG berikut kedalam NLTK CFG fromstring:

Kalimat → Subyek Predikat

Kalimat → Subyek Predikat Keterangan

Subyek → Determinan KataBenda

Predikat → KataKerja KataBenda

Determinan → seorang, seekor

KataBenda → manusia, anjing, nasi, daging, air

KataKerja → makan, minum

Keterangan → kemaren, tadi

- lakukan percobaan dengan kalimat berikut

“seekor anjing makan daging tadi” dan “seorang manusia minum air kemaren”

- Tampilkan tree-nya keduanya



# Prosedur Pengumpulan Tugas

- Buatlah laporan dalam slide ppt. laporan terdiri dari screenshot coding, hasil running dan analisa untuk setiap tugas.
- Upload file tersebut ke ETHOL
- Simpan laporan dalam file pdf dengan format penamaan:
  - **Untuk D4 IT A:**  
AID4ITA\_M6\_NRP\_namadepan.pdf  
Deadline upload: Senin, 1 April Pukul 23.59 WIB
  - **Untuk D3 IT A:**  
AID3ITA\_M6\_NRP\_namadepan.pdf  
Deadline upload: Senin, 1 April Pukul 23.59 WIB



# Referensi

- Modul Ajar Kecerdasan Buatan, Entin Martiana, Ali Ridho Barakbah, Yuliana Setiowati, Politeknik Elektronika Negeri Surabaya, 2014.
- Artificial Intelligence (Teori dan Aplikasinya), Sri Kusumadewi, cetakan pertama, Penerbit Graha Ilmu, 2003.



**bridge to the future**

<http://www.eepis-its.edu>