

# Teknologi Kecerdasan Artifisial

## Kuliah Team Teaching Pekan 10

### Natural Language Processing



Team Teaching Universitas Gunadarma  
@ 2023 :

1. Dr. Mohammad Iqbal
2. Dr. Nur Sultan Salahuddin
3. Dr. Raden Supriyanto
4. Dr. Sunny Arief
5. Ike Putri Kusumawijaya, ST, MMSI



# Agenda



**1**

Memahami NLP

**2**

Komponen dan  
Proses dalam NLP

**3**

Contoh aplikasi NLP  
dalam kehidupan

---

# Memahami Natural Language Processing

---

# Apa itu NLP? (1)

- **Natural Language Processing (NLP)** adalah bidang Artificial Intelligence (AI) yang membuat bahasa manusia dapat dipahami oleh mesin. NLP menggabungkan kekuatan **linguistik** dan **ilmu komputer** untuk mempelajari aturan dan struktur bahasa, dan menciptakan **sistem cerdas** yang mampu memahami, menganalisis, dan mengekstraksi makna dari teks dan ucapan.

Sumber : <https://monkeylearn.com/natural-language-processing>

# Apa itu NLP? (2)

- Natural Language Processing (NLP) adalah cabang **kecerdasan buatan** yang membantu komputer **memahami, menafsirkan, dan memanipulasi bahasa manusia**.
- NLP diambil dari **banyak disiplin ilmu**, termasuk ilmu komputer dan linguistik komputasi, dalam upayanya untuk mengisi **kesenjangan** antara komunikasi manusia dan pemahaman komputer.

# Apa itu NLP? (3)

- Pemrosesan bahasa natural difokuskan pada desain dan analisis algoritma komputasi dan representasi untuk memproses bahasa alami manusia [Eisenstein, 2018]
- NLP adalah bidang yang mempelajari *metode, algoritma, representasi* yang memproses input ataupun output yang berupa Bahasa alami, baik berupa teks maupun suara.



# Apa itu NLP? (4)

- Tujuan utama NLP adalah untuk membaca, menguraikan, dan memahami bahasa manusia.
- Sebagian besar teknik NLP mengandalkan *machine learning* untuk mendapatkan makna dari bahasa manusia.

# Untuk Apa NLP Digunakan?



Memahami struktur dan makna bahasa manusia dengan menganalisis berbagai aspek seperti **sintaksis, semantik, pragmatik, dan morfologi**. Kemudian, ilmu komputer mengubah pengetahuan linguistik ini menjadi algoritma *machine learning* berbasis aturan yang dapat memecahkan masalah tertentu dan melakukan tugas yang diinginkan.



# Mengapa NLP ? (1)

## *Volume data tekstual yang besar*

- Pemrosesan bahasa alami membantu komputer berkomunikasi dengan manusia dalam bahasa mereka sendiri dan menskalakan tugas terkait bahasa lainnya. Misalnya, NLP memungkinkan komputer membaca teks, mendengar ucapan, menafsirkannya, mengukur sentimen, dan menentukan bagian mana yang penting.
- Mesin saat ini dapat menganalisis lebih banyak data berbasis bahasa daripada manusia, tanpa kelelahan dan dengan cara yang konsisten dan tidak bias. Mengingat banyaknya data tidak terstruktur yang dihasilkan setiap hari, dari rekam medis hingga media sosial, otomatisasi sangat penting untuk menganalisis data teks dan ucapan secara efisien.

# Mengapa NLP ? (2)

## Menyusun data yang sangat tidak terstruktur

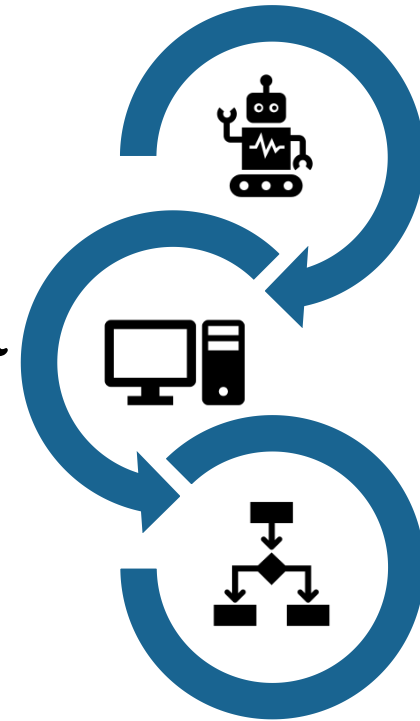
- Bahasa manusia sangat kompleks dan beragam. Manusia mengekspresikan dirinya dengan cara yang tak terbatas, baik secara lisan maupun tertulis. Tidak hanya terdapat ratusan bahasa dan dialek, tetapi dalam setiap bahasa terdapat seperangkat aturan tata bahasa dan sintaksis, istilah, dan bahasa gaul yang unik. Saat menulis, manusia sering salah mengeja atau menyingkat kata, atau menghilangkan tanda baca.

# Mengapa NLP ? (3)

- Meskipun pembelajaran yang diawasi dan tidak diawasi, dan khususnya *deep learning*, sekarang banyak digunakan untuk pemodelan bahasa manusia, ada juga kebutuhan untuk pemahaman sintaksis dan semantik serta keahlian domain yang tidak selalu ada dalam pendekatan *machine learning*. NLP penting karena membantu menyelesaikan ambiguitas dalam bahasa dan menambahkan struktur numerik yang berguna ke data untuk banyak aplikasi, seperti pengenalan ucapan atau analitik teks

# Mengapa NLP? (4)

- *Teknologi yang human friendly*
- *Peningkatan daya komputasi*
- *Algoritma yang terus ditingkatkan*



# Disiplin Ilmu Dari NLP



# Fonetik / Fonologi



Berhubungan dengan suara yang menghasilkan kata yang dapat dikenali. Bidang ini dipakai dalam aplikasi-aplikasi *speech based system*




# Morfologi

Pengetahuan  
tentang *kata* dan  
*bentuknya* sehingga  
bisa dibedakan  
antara yang satu  
dengan yang lain



# Sintaksis

Pengetahuan  
tentang urutan  
kata dalam  
pembentukan  
kalimat



Kalimat  $\rightarrow S, P$   
 $S \rightarrow$  Determinan,  
kata benda  
 $P \rightarrow$  Kata Kerja,  
Kata benda

# Semantik

Mempelajari arti  
suatu kata dan  
bagaimana arti  
kata-kata  
tersebut  
membentuk suatu  
arti kata dari  
kalimat yang utuh




# Pragmatik

- Pengetahuan tentang konteks kata/kalimat yang berhubungan erat dengan keadaan atau situasi kata/kalimat tersebut dipakai
- Contoh:
  - Paman pulang (diucapkan dengan nada datar)
  - Paman pulang! (diucapkan dengan nada tinggi)
  - Paman pulang? (diucapkan dengan nada tanya)



# 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 di sana.

# World Knowledge



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



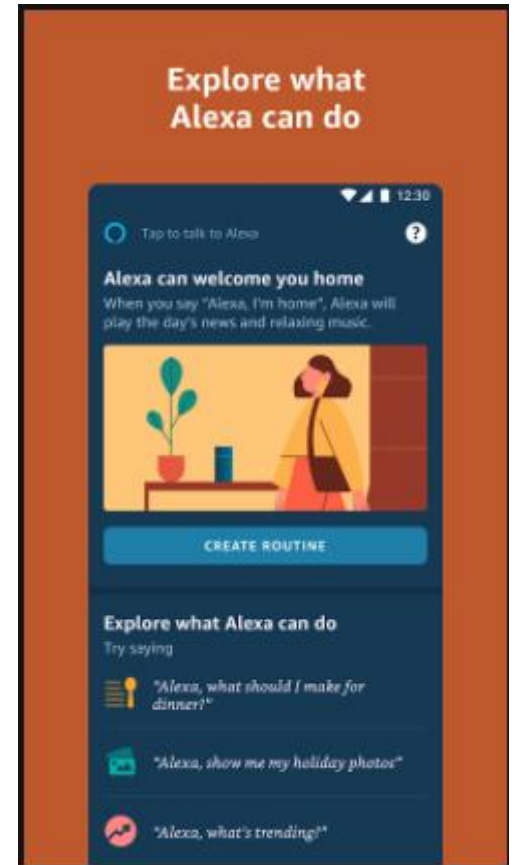


# Evolusi Pemrosesan Bahasa Alami (1)

- Meskipun pemrosesan bahasa alami bukanlah ilmu baru, teknologinya berkembang pesat berkat peningkatan minat dalam komunikasi manusia-ke-mesin, ditambah ketersediaan data besar, komputasi yang andal, dan algoritma yang ditingkatkan.
- Manusia dapat berbicara dan menulis dalam bahasa Inggris, Spanyol, atau China. Tetapi bahasa asli komputer – yang dikenal sebagai kode mesin atau bahasa mesin – sebagian besar tidak dapat dipahami oleh kebanyakan orang. Pada tingkat perangkat yang paling rendah, komunikasi terjadi bukan dengan kata-kata tetapi melalui jutaan angka nol dan yang menghasilkan

# Evolusi Pemrosesan Bahasa Alami (2)

70 tahun yang lalu, pemrogram menggunakan kartu berlubang untuk berkomunikasi dengan komputer pertama. Proses manual dan sulit ini dipahami oleh sejumlah kecil orang. Sekarang Anda dapat mengatakan, "Alexa, saya suka lagu ini," dan perangkat yang memutar musik di rumah Anda akan menurunkan volume dan menjawab, "Oke. Peringkat disimpan," dengan suara yang mirip manusia. Kemudian ia menyesuaikan algoritmanya untuk memutar lagu itu - dan lagu lainnya yang disukai saat Anda



# Evolusi Pemrosesan Bahasa Alami (3)

*Mengurangi keluhan pelanggan dengan NLP:*

Royal Bank of Scotland menggunakan analitik teks, teknik NLP, untuk mengekstrak tren penting dari umpan balik pelanggan dalam berbagai bentuk.

Perusahaan menganalisis data dari email, survei, dan percakapan call center untuk mengidentifikasi akar penyebab ketidakpuasan pelanggan dan menerapkan perbaikan.

# Mengapa NLP Sulit? (1)

- Sifat bahasa manusia yang membuat NLP sulit.
- Aturan yang mengatur penyampaian informasi menggunakan bahasa alami tidak mudah dipahami oleh komputer.
- Beberapa dari aturan ini dapat bersifat tingkat tinggi dan abstrak; Misalnya, ketika seseorang menggunakan ucapan sarkastik untuk menyampaikan informasi.
- Di sisi lain, aturan untuk menyatakan sesuatu bersifat pluralitas item misalnya, menggunakan karakter "s"

# Mengapa NLP Sulit? (2)

- Memahami bahasa manusia secara komprehensif membutuhkan pemahaman baik kata-kata maupun bagaimana konsep-konsep tersebut dihubungkan untuk menyampaikan pesan yang dimaksudkan.
- Meskipun manusia dapat dengan mudah menguasai suatu bahasa, ambiguitas dan karakteristik yang tidak tepat dari bahasa alami inilah yang membuat NLP sulit untuk diimplementasikan oleh mesin.

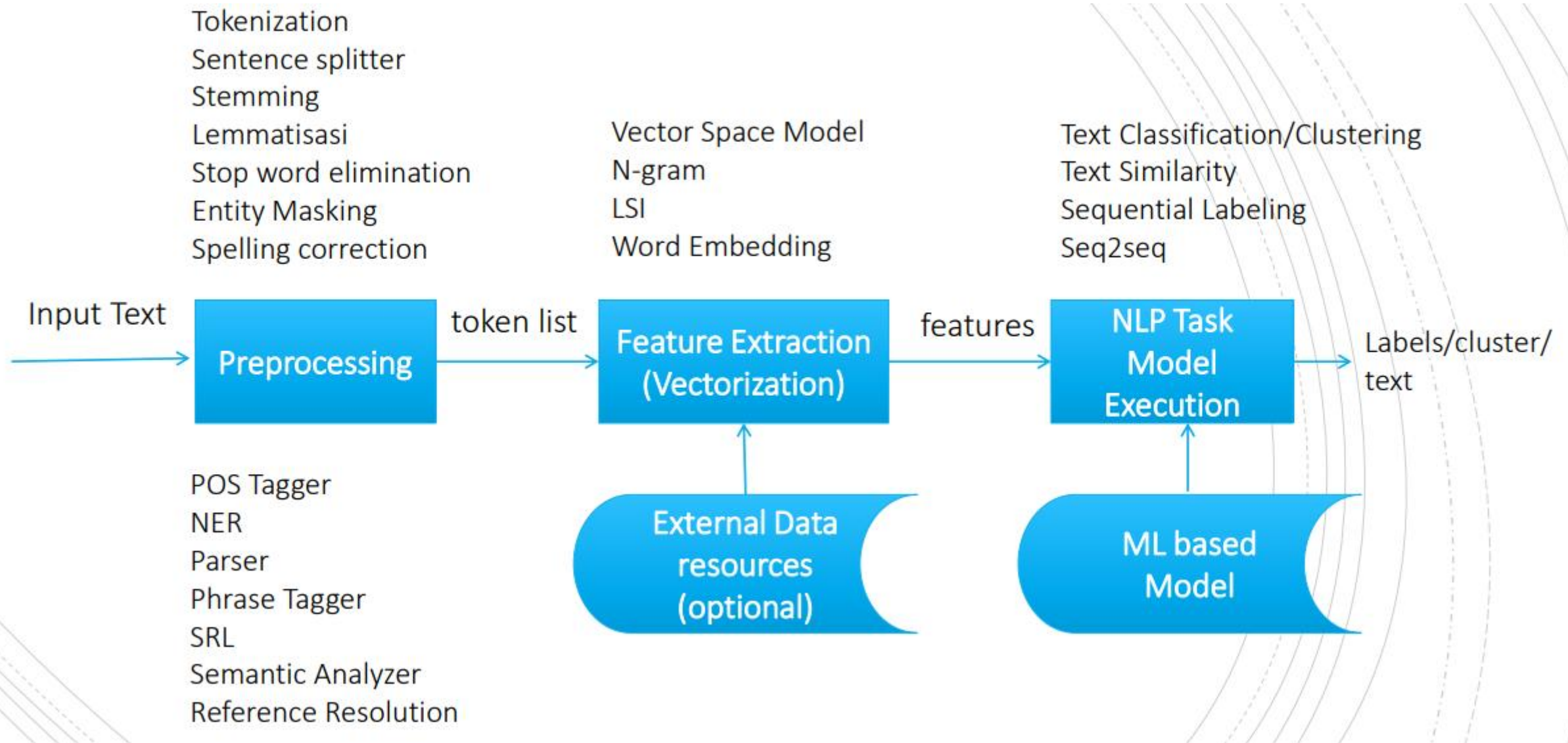
---

# Komponen dan Proses dalam NLP

---



# Kerangka Kerja dari NLP



# Bagaimana cara kerja NLP? (1)

## Memecah potongan-potongan unsur Bahasa

- Pemrosesan bahasa alami mencakup banyak teknik berbeda untuk menafsirkan bahasa manusia, mulai dari metode statistik dan *machine learning* hingga pendekatan berbasis aturan dan algoritmik. Pendekatan yang luas dibutuhkan karena data berbasis teks dan suara sangat bervariasi, seperti halnya aplikasi praktis.
- Tugas NLP dasar termasuk tokenisasi dan parsing, lemmatization / stemming, penandaan part-of-speech, deteksi bahasa dan identifikasi hubungan semantik.

# Bagaimana cara kerja NLP? (2)

- Secara umum, tugas NLP memecah bahasa menjadi bagian-bagian yang lebih pendek dan mendasar, mencoba memahami hubungan antara bagian-bagian tersebut dan mengeksplorasi bagaimana bagian-bagian tersebut bekerja sama untuk menciptakan makna.

# Teknik yang Digunakan dalam NLP (1)

Analisis sintaksis dan analisis semantik adalah teknik utama yang digunakan untuk menyelesaikan tugas Pemrosesan Bahasa Alami. Berikut ini penjelasan tentang bagaimana mereka dapat digunakan.

## 1. Sintaks

Sintaksis mengacu pada susunan kata-kata dalam sebuah kalimat sehingga membuat pengertian gramatikal. Dalam NLP, analisis sintaksis digunakan untuk menilai bagaimana bahasa alami selaras dengan aturan tata bahasa. Algoritme komputer digunakan untuk menerapkan aturan tata bahasa ke sekelompok kata dan mendapatkan maknanya.

# Teknik yang Digunakan dalam NLP (2)

Berikut beberapa teknik sintaks yang dapat digunakan:

- Lemmatisasi: Ini memerlukan pengurangan berbagai bentuk kata yang berubah menjadi satu bentuk untuk memudahkan analisis.
- Segmentasi morfologis: Ini melibatkan pembagian kata menjadi unit-unit individu yang disebut morfem.
- Segmentasi kata: Ini melibatkan pembagian sebagian besar teks berkelanjutan menjadi unit-unit yang berbeda.

# Teknik yang Digunakan dalam NLP (3)

- Parts-of-Speech: Ini melibatkan pengidentifikasian bagian ucapan untuk setiap kata.
- Parsing: Ini melibatkan melakukan analisis tata bahasa untuk kalimat yang disediakan.
- Pemutusan kalimat: Ini melibatkan penempatan batas kalimat pada sepotong teks besar.
- Stemming: Ini melibatkan pemotongan kata-kata yang diubah ke bentuk akarnya.



# Teknik yang Digunakan dalam NLP (4)

## 2. Semantik

Semantik mengacu pada makna yang disampaikan oleh sebuah teks. Analisis semantik adalah salah satu aspek sulit dari Natural Language Processing yang belum sepenuhnya terselesaikan.

Ini melibatkan penerapan algoritma komputer untuk memahami arti dan interpretasi kata dan bagaimana kalimat disusun.

# Teknik yang Digunakan dalam NLP (5)

## 2. Semantik (lanjutan)

Manusia berkomunikasi berdasarkan makna dan konteks. Semantik membantu komputer mengidentifikasi struktur kalimat dan elemen teks yang paling relevan untuk memahami topik yang sedang dibahas. Misalnya, jika teks berisi kata-kata seperti pemilu, demokrat dan republik atau anggaran, pajak dan inflasi, komputer memahami bahwa topik yang dibahas adalah politik dan ekonomi.

# Teknik yang Digunakan dalam NLP (6)

Berikut beberapa teknik dalam analisis semantik:

- Name Entity Recognition (NER): Ini melibatkan penentuan bagian-bagian teks yang dapat diidentifikasi dan dikategorikan ke dalam grup preset. Contoh kelompok seperti itu termasuk nama orang dan nama tempat.
- Disambiguasi arti kata: Ini melibatkan pemberian makna pada kata berdasarkan konteksnya.
- Generasi bahasa alami: Ini melibatkan penggunaan database untuk mendapatkan maksud semantik dan mengubahnya menjadi bahasa manusia.

# Memahami cara membangun NLP (1)

- Perhatikan paragraf berikut:

*London is the capital and most populous city of England and the United Kingdom. Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded by the Romans, who named it Londinium.*

- Paragraf ini berisi beberapa fakta berguna. Alangkah baiknya jika komputer dapat membaca teks ini dan memahami bahwa **London** adalah sebuah **kota**, **London** terletak di **Inggris**, **London** dihuni oleh orang **Romawi** dan seterusnya. Tetapi untuk mencapainya, pertama-tama kita harus mengajari komputer konsep yang paling dasar dari Bahasa tertulis

# Memahami cara membangun NLP (2)

Langkah-Langkah untuk membangun NLP

## 1. Segmentasi Kalimat

Langkah pertama adalah **memecah teks** menjadi beberapa kalimat terpisah. Proses ini menghasilkan kalimat sebagai berikut:

1. London is the capital and most populous city of England and the United Kingdom.
2. Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia.
3. It was founded by the Romans, who named it Londinium.

# Memahami cara membangun NLP (3)

## 2. Tokenisasi Kata

Setelah membagi dokumen menjadi beberapa kalimat, lalu **memecah kalimat menjadi kata atau token**.

Contoh:

*"London is the capital and most populous city of England and the United Kingdom."*

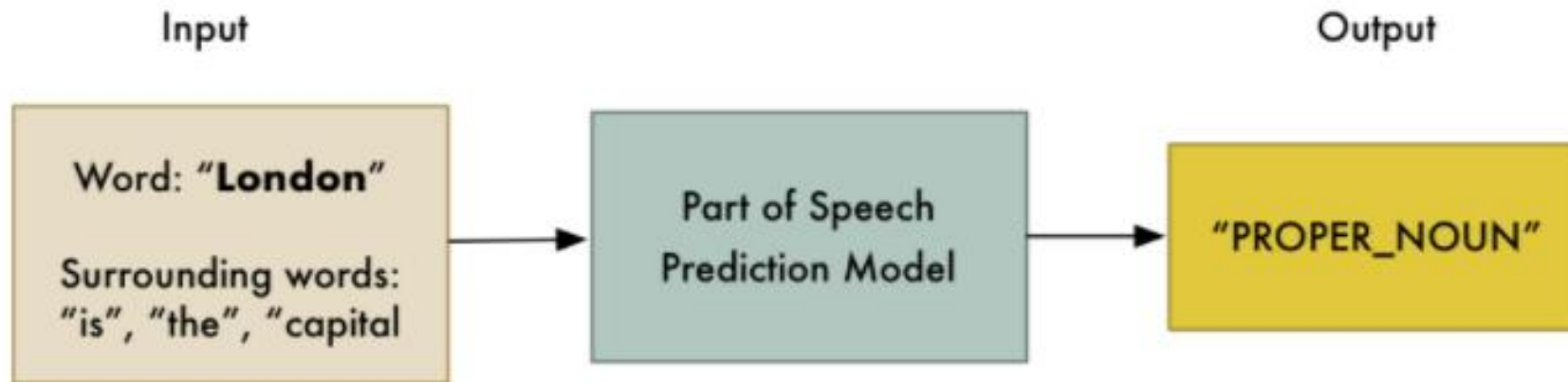


*"London", "is", "the", "capital", "and", "most",  
"populous", "city", "of", "England", "and", "the",  
"United", "Kingdom", "."*

# Memahami cara membangun NLP (4)

## 3. Memprediksi Bagian Ucapan (Part-of-Speech) untuk Setiap Token

Selanjutnya akan dilihat setiap token atau kata dan mencoba menentukan - apakah itu **kata benda**, **kata kerja**, **kata sifat**, dan sebagainya. Mengetahui peran setiap kata dalam kalimat akan membantu memahami



# Memahami cara membangun NLP (5)

Dengan informasi ini, mulai terkumpul beberapa makna yang sangat mendasar. Misalnya, bahwa kata benda dalam kalimat tersebut menyertakan “London” dan “capital”, jadi kalimat tersebut mungkin berbicara tentang London.

## 4. Lemmatisasi Teks

Dalam bahasa Inggris (dan sebagian besar bahasa), kata-kata muncul dalam berbagai bentuk. Lihat dua kalimat ini:

I had a pony.

I had two ponies.



# Memahami cara membangun NLP (6)

Kedua kalimat tersebut berbicara tentang kata benda **pony**, tetapi menggunakan infleksi yang berbeda. Saat bekerja dengan teks di komputer, akan sangat membantu jika mengetahui bentuk dasar setiap kata sehingga tahu bahwa kedua kalimat tersebut berbicara tentang konsep yang sama. Jika tidak, string "pony" dan "ponies" terlihat seperti dua kata yang sama sekali berbeda di komputer.

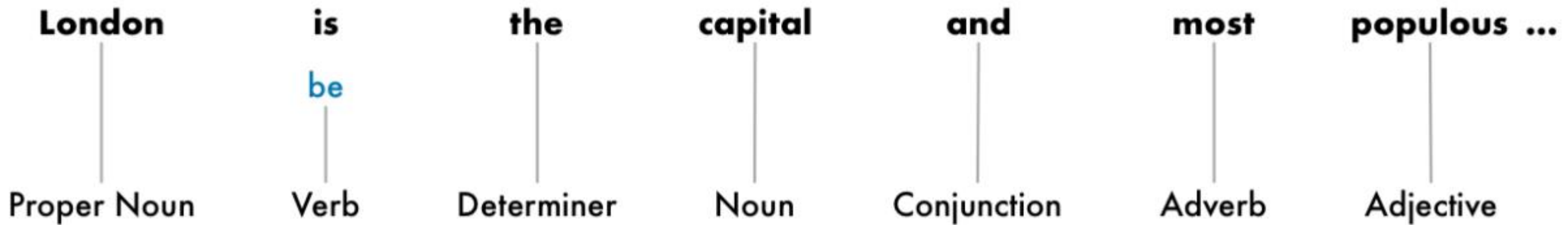
Pada NLP disebut lemmatization – mencari tahu bentuk atau lemma paling dasar dari setiap kata dalam kalimat.

# Memahami cara membangun NLP (7)

- Hal yang sama berlaku untuk kata kerja. Kita juga bisa membuat lemmatisasi kata kerja dengan mencari akarnya, bentuk tak terkonjugasi. Jadi, " *I had two ponies* " menjadi " *I [have] two [pony]* ."
- Lemmatisasi biasanya dilakukan dengan memiliki tabel pencarian bentuk kata lemma berdasarkan bagian ucapannya dan mungkin memiliki beberapa aturan khusus untuk menangani kata yang belum pernah dilihat sebelumnya.

# Memahami cara membangun NLP (8)

- Berikut tampilan kalimat pada contoh setelah lemmatisasi ditambahkan dalam bentuk akar kata kerja:

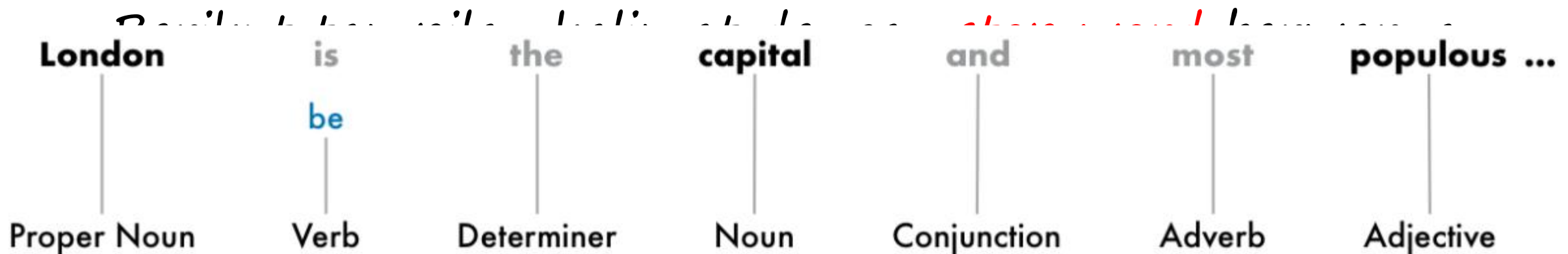


- Satu-satunya perubahan yang terjadi adalah mengubah "*is*" menjadi "*be*".

# Memahami cara membangun NLP (9)

## 5. Mengidentifikasi Stop Word

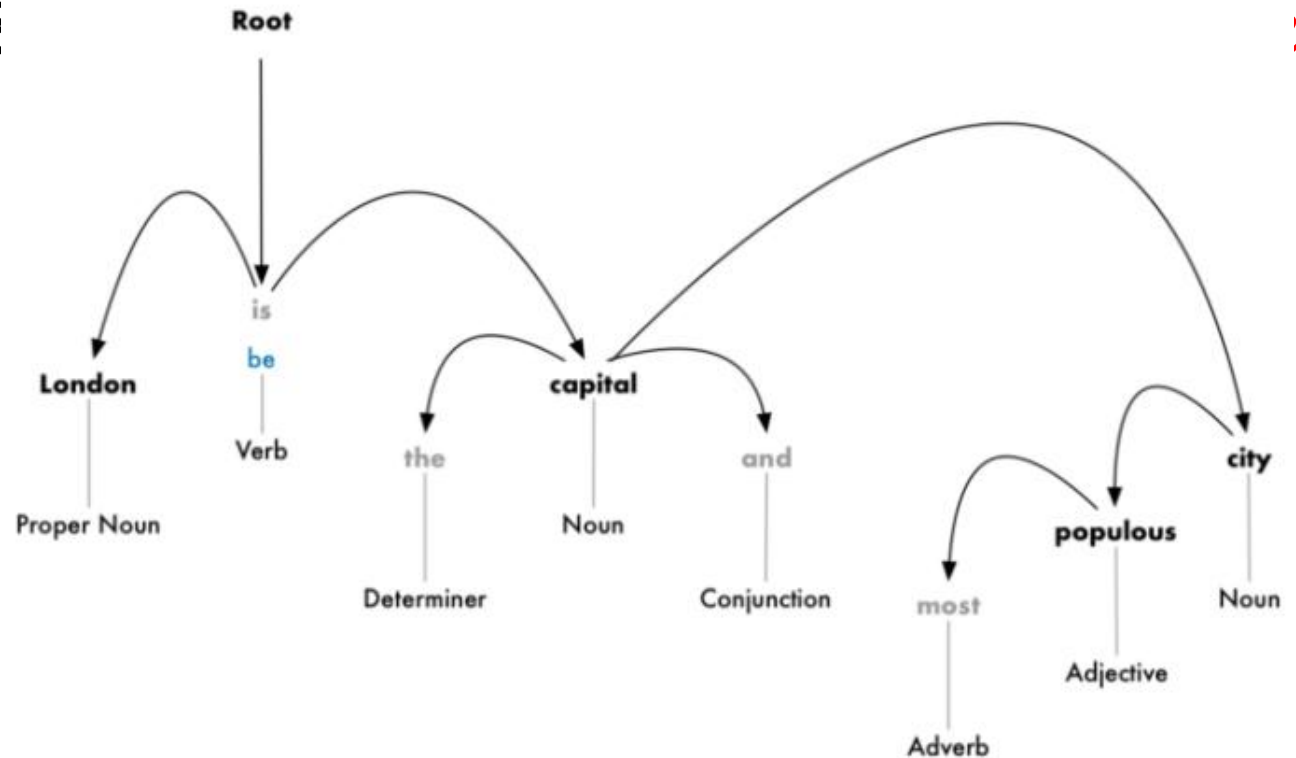
Bahasa Inggris memiliki banyak kata pengisi yang sering muncul seperti "and", "the", dan "a". Saat melakukan statistik pada teks, kata-kata ini menimbulkan banyak gangguan karena lebih sering muncul daripada kata lain yang **lebih penting**. Beberapa NLP akan menandainya sebagai **stop words** —yaitu, kata yang mungkin ingin difilter sebelum melakukan analisis statistik.



# Memahami cara membangun NLP (10)

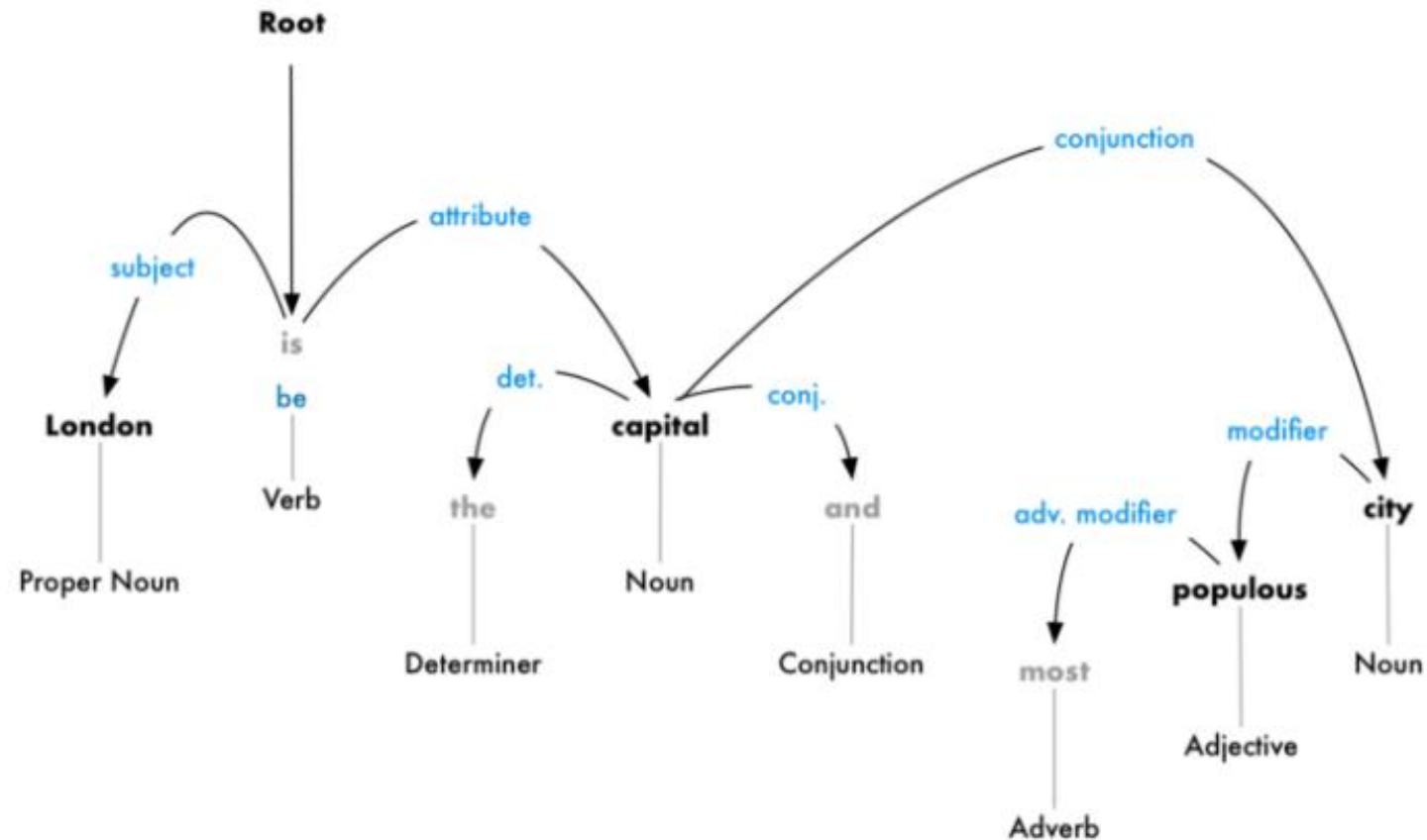
## 6. Dependency Parsing

Tujuannya adalah untuk membangun pohon yang menetapkan satu kata induk untuk setiap kata dalam kalimat. Akar pohon akan menjadi kata kerja utama dalam kalimat. E



# Memahami cara membangun NLP (11)

*Selain mengidentifikasi kata induk dari setiap kata, juga dapat memprediksi jenis hubungan yang ada di antara kedua kata tersebut.*



# Memahami cara membangun NLP (12)

- Pohon parse ini menunjukkan bahwa subjek kalimat adalah kata benda "London" dan memiliki hubungan "be" dengan "capitol". Kita akhirnya tahu sesuatu yang berguna – *London is a capital!* Dan jika kita mengikuti pohon parse lengkap untuk kalimat, kita bahkan akan menemukan bahwa *London is the capital of the United Kingdom.*

# Memahami cara membangun NLP (13)

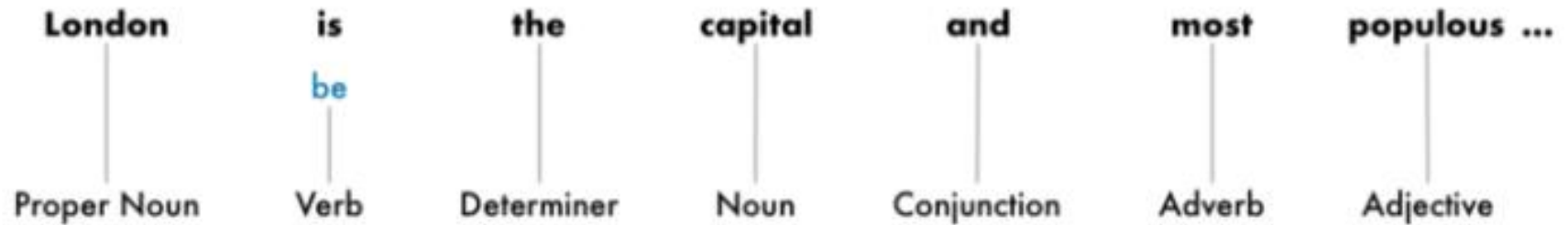
## 7. Menemukan Naoun Phrases (Frasa Kata Benda)

Sejauh ini, setiap kata dalam kalimat diperlakukan sebagai entitas terpisah. Namun terkadang lebih masuk akal untuk mengelompokkan kata-kata yang mewakili satu ide. Kita dapat menggunakan informasi dari pohon penguraian ketergantungan untuk secara otomatis mengelompokkan kata-kata yang semuanya membicarakan hal yang sama.



# Memahami cara membangun NLP (14)

Misal :



Kita dapat mengelompokkan frase kata benda untuk menghasilkan ini:



# Memahami cara membangun NLP (15)

## 8. Named Entity Recognition (NER)

Hasil dari Langkah 7 terdapat kata benda berikut:

**London** is the **capital** and most populous **city** of **England** and the **United Kingdom**.

Beberapa dari kata benda ini menyajikan hal-hal nyata di dunia. Misalnya, "London", "England", dan "United Kingdom" mewakili tempat fisik di peta. Alangkah baiknya bisa mendeteksinya! Dengan informasi tersebut, dapat secara otomatis mengekstrak daftar tempat dunia nyata yang disebutkan dalam dokumen menggunakan NLP.

# Memahami cara membangun NLP (16)

- Tujuan dari Named Entity Recognition, atau NER, adalah untuk mendeteksi dan memberi label kata benda ini dengan konsep dunia nyata yang mereka wakili. Berikut tampilan kalimat setelah menjalankan setiap token melalui model pemberian tag NER:

**London** is the capital and most populous city of **England** and the **United Kingdom**.

Geographic  
Entity

Geographic  
Entity

Geographic  
Entity

# Memahami cara membangun NLP (17)

- Tetapi sistem NER tidak hanya melakukan pencarian kamus sederhana. Sebaliknya, mereka menggunakan konteks bagaimana sebuah kata muncul dalam kalimat dan model statistik untuk menebak jenis kata benda yang diwakili oleh sebuah kata. Sistem NER yang baik dapat membedakan antara *orang* "Brooklyn Decker" dan *tempat* "Brooklyn" menggunakan petunjuk konteks.

# Memahami cara membangun NLP (18)

Jenis objek yang dapat  
diberi tag oleh sistem NER

**Nama Orang**

**Nama Perusahaan**

**Lokasi Geografis**

**Nama Produk**

**Waktu / Jam dan tanggal**

**Sejumlah uang**

**Nama Peristiwa**

# Memahami cara membangun NLP (19)

## 9. Coreference Resolution

Pada titik ini, kita sudah memiliki representasi yang berguna dari kalimat kita. Kita tahu bagian-bagian kata untuk setiap kata, bagaimana kata-kata itu berhubungan satu sama lain, dan kata-kata mana yang berbicara tentang entitas bernama.

Namun, kita masih memiliki satu masalah besar. Bahasa Inggris penuh dengan kata ganti seperti he, she, dan it. Ini adalah yang biasa digunakan daripada menuliskan nama berulang kali di setiap kalimat. Manusia dapat melacak apa yang diwakili oleh kata-kata ini berdasarkan konteksnya. Namun model NLP tidak mengetahui arti kata ganti karena hanya memeriksa

# Memahami cara membangun NLP (20)

- Mari kita lihat kalimat ketiga dalam dokumen:

*"It was founded by the Romans, who named it Londinium."*

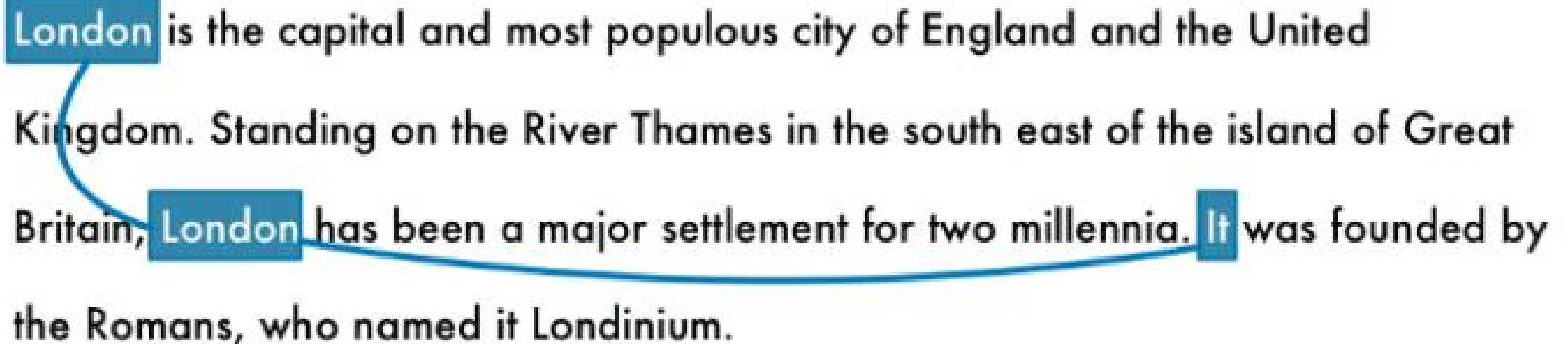
Jika kita mengurai ini dengan NLP, kita akan tahu bahwa "it" didirikan oleh orang Romawi. Tetapi jauh lebih berguna untuk mengetahui bahwa "London" didirikan oleh orang Romawi.

Sebagai manusia yang membaca kalimat ini, dapat dengan mudah mengetahui bahwa "it" berarti "London". Tujuan dari resolusi inti adalah untuk menemukan pemetaan yang sama ini dengan melacak kata ganti di seluruh kalimat

# Memahami cara membangun NLP (21)

- Berikut adalah hasil dari menjalankan Coreference Resolution pada dokumen untuk kata "London":

London is the capital and most populous city of England and the United Kingdom. Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded by the Romans, who named it Londinium.

A blue line connects the word "London" in the first sentence to the word "London" in the third sentence, indicating they refer to the same entity. Another blue line connects the word "It" in the third sentence to the word "London" in the same sentence, indicating a coreference between the pronoun and the noun.



# Alur Proses Penyaringan Spam Berbasis Aturan Manual

- Data → Feature
  - **Text**: “complimentary Ibiza Holiday needs your URGENT collection”
  - **Token list**: complimentary, ibiza, holiday, needs, your, urgent, collection
  - **Spam word list**: urgent, holiday
- Prosedur → Teknik Klasifikasi
  - Setiap kata dalam daftar token diperiksa ke dalam daftar kata spam
  - Hitung berapa banyak kata dalam daftar token yang merupakan kata-kata spam
  - Teks diberi label sebagai spam jika jumlah kata spam

---

# Contoh Aplikasi NLP dalam Kehidupan sehari-hari

---

# Beberapa Aplikasi NLP

- 1 Information Retrieval
- 2 Information Extraction
- 3 Question Answering
- 4 Text Summarization
- 5 Machine Translation
- 6 Lain-lain

# Information Retrieval

- Query: “daftar restoran sushi Jakarta”
- Mesin pencari menggunakan pemrosesan bahasa alami untuk menghasilkan hasil pencarian yang relevan berdasarkan perilaku pencarian atau maksud pengguna yang serupa.

Peringkat ▾

Harga ▾

Jam buka ▾


Sushi Tengoku

4,5 ★★★★★ (2.055) · \$S · Sushi

Jl. Radio Dalam Raya No.11 B, RT.5/RW.4

Segera tutup · Tutup pukul 23.00

✓ Makan di tempat · ✓ Bawa pulang · ✓ Pesan antar




Sushi Hiro Pondok Indah Mall

4,2 ★★★★★ (29) · Sushi

Pd. Indah Mall Street Gallery, Jl. Metro Pondok Indah No.106, RT.1/RW.16

Segera tutup · Tutup pukul 23.00

✓ Makan di tempat · ✓ Bawa pulang · ✓ Antar tanpa bertemu




Sushi-Ya

4,2 ★★★★★ (276) · \$S · Sushi

Jl. Tebet Raya No.53A, RT.3/RW.2

✓ Makan di tempat · ✓ Bawa pulang · ✓ Antar tanpa bertemu



→

Lihat semua

<https://www.cosmopolitan.co.id> > article > read > daftar... ▾

## Daftar Restoran Sushi Terbaik di Jakarta

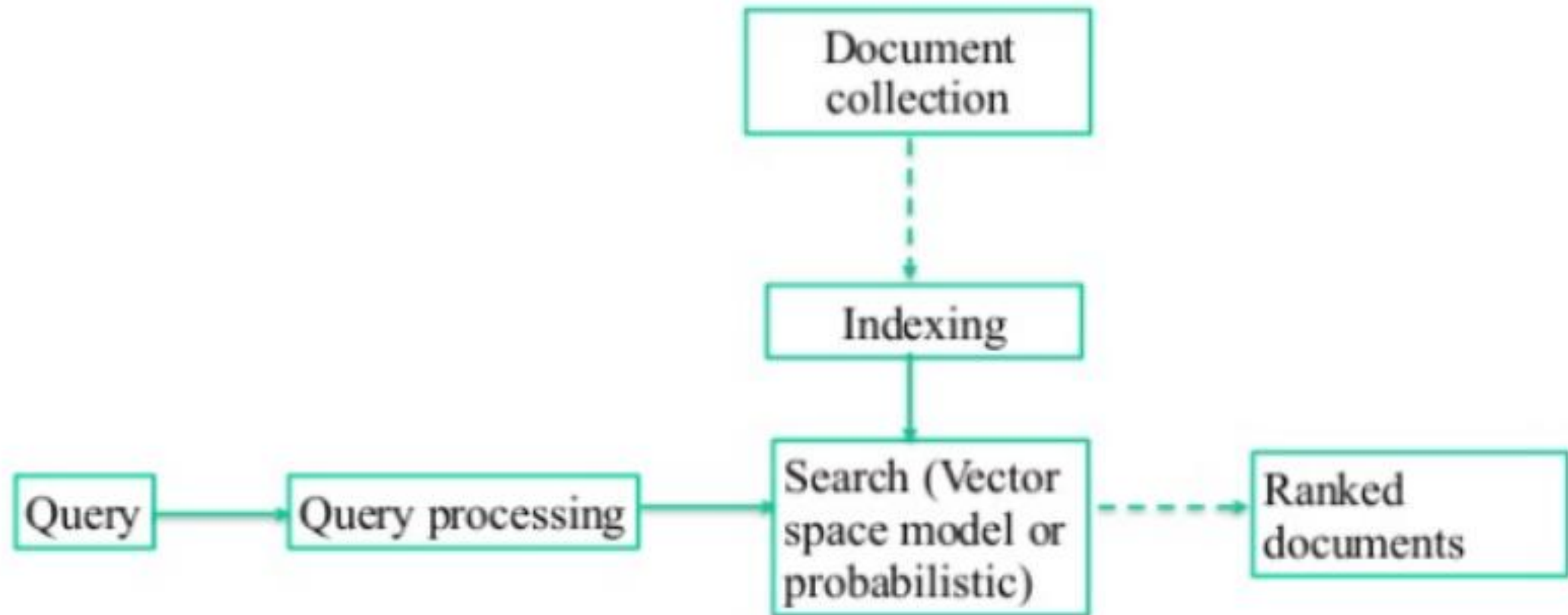
27 Mei 2018 — Sushi lovers, apakah kamu sudah mencicipi sushi dan sashimi dari 5 restoran Jepang berikut? Inilah daftar restoran yang menyajikan sushi ...

<https://travel.kompas.com> > Travel > Makan Makan ▾

## Sushi Tei sampai Nama Sushi, Daftar 11 Restoran di Jakarta ...

Sushi Tei sampai Nama Sushi, Daftar 11 Restoran di Jakarta dengan Layanan Pesan Antar Sushi. Kompas.com - 27/04/2020, 12:11 WIB. Bagikan: Komentar.

# Arsitektur Sistem Information Retrieval



# Information Extraction

- Untuk mengekstrak dari teks tidak terstruktur, informasi yang telah ditentukan sebelumnya atau ditentukan sebelumnya dalam template – Isi sejumlah slot / atribut
- Contoh: gunakan template

[PERSON, go, LOCATION, TIME]

untuk mengekstrak informasi tentang tujuan individu pergi.

“President Obama went to Hanoi yesterday.

[PERSON = "Presiden Obama", go, LOCATION = "Hanoi", TIME = "yesterday"]

# Question Answering (1)

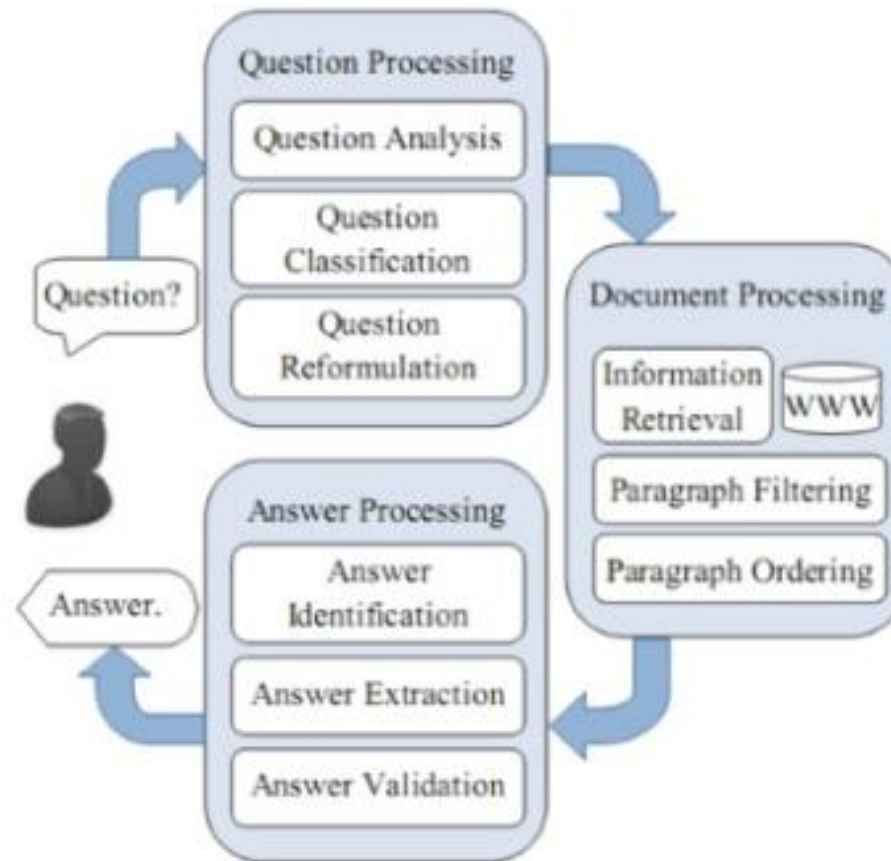
- Sistem yang secara otomatis mengembalikan jawaban atas pertanyaan pengguna dengan mengambil informasi dari dokumen yang dikumpulkan.
- Perbedaan dengan Information retrieval:
  - Sasaran sistem QA adalah untuk menanggapi jawaban yang tepat, bukan dokumen yang terkait dengan pertanyaan pengguna.
    - T: siapa yang menciptakan internet? J: Robert E. Kahn dan Vint Cerf.
  - Sistem QA membutuhkan analisis semantik yang lebih rumit.

# Question Answering (2)

- Jawaban pertanyaan factoid:
  - Siapa / Apa / Dimana / Kapan
  - Jawaban seringkali berupa frase.
- Jawaban pertanyaan non-factoid:
  - Pertanyaan definisi
  - Bagaimana / Mengapa
  - Jawaban dapat mencakup banyak kalimat (paragraf)

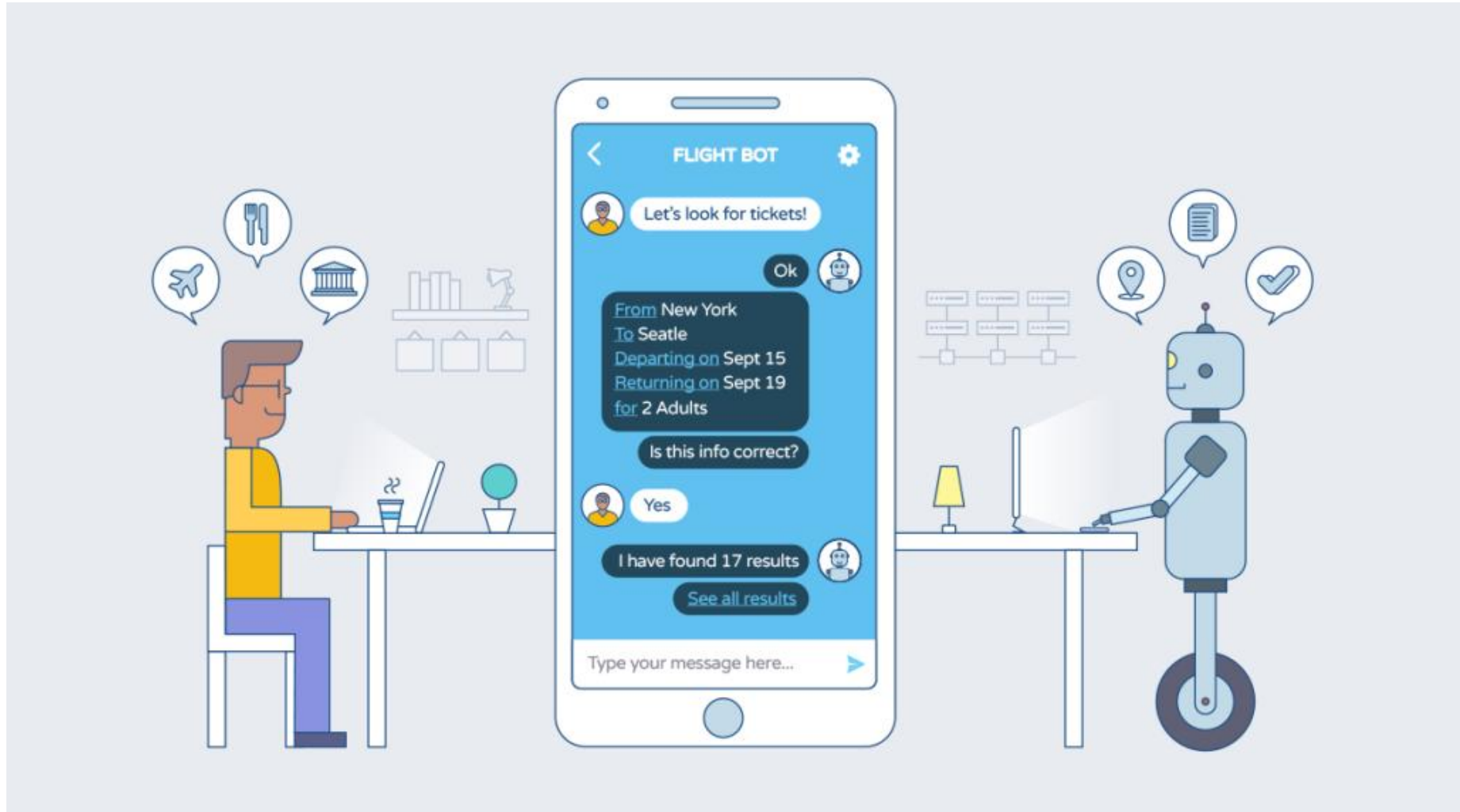


# Question Answering (3)



Sumber: Dr Ngo Xuan Bach: <http://tinyurl.com/jk2dv33>

# Contoh QA : Chatbot (1)



## Contoh QA : Chatbot (2)

- Chatbot merupakan layanan digital yang mampu melakukan percakapan di sebuah aplikasi pesan instan. Percakapan yang dilakukan terlihat natural, seperti melakukan percakapan dengan manusia. Hal ini bisa terjadi karena adanya artificial intelligence atau kecerdasan buatan.
- Dalam menjalankan bisnis, pasti harus membalas pertanyaan dari konsumen. Bila manusia memegang chat sendiri, pasti akan kewalahan. Disinilah kegunaan chatbot, karena bisa membalas pertanyaan konsumen lebih cepat.

# Kelebihan Chatbot (1)

- **Pelayanan Bisnis 24 Jam.** Kelebihan chatbot adalah anda bisa memberikan pelayanan konsumen 24 jam penuh. Karena sistem ini tidak terpengaruh oleh jam kerja customer service, Yang memiliki jam kerja dari pukul 8 pagi sampai 5 sore.
- **Respon Menjawab Pertanyaan Konsumen Lebih Cepat.** Konsumen pasti ingin pertanyaannya cepat di jawab. Kalau memakai customer service manual, terkadang CS sibuk melakukan berbagai hal dan akhirnya responnya lambat. Membuat konsumen pergi, tetapi dengan memakai chatbot respon menjawab pertanyaan konsumen bisa lebih cepat.

# Kelebihan Chatbot (2)

- **Interaksi Dengan Konsumen Lebih Mudah.** Membuat keterikatan terjalin dengan baik. Inilah salah satu kelebihan chatbot. Konsumen akan lebih nyaman berbelanja. Karena semua pertanyaan mereka tentang produk, akan dijawab dengan cepat. Serasa membeli langsung di toko offline. Seperti berbincang dengan pemilik tokonya langsung.
- **Respon Terhadap Pertanyaan dan Keluhan Kompleks.** Terkadang admin malas untuk menjawab pertanyaan konsumen yang terlalu kompleks atau keluhan yang membingungkan. Manusia tentunya memiliki rasa malas, emosi dan lainnya. Tetapi berbeda dengan chatbot. Chatbot mampu merespon pertanyaan

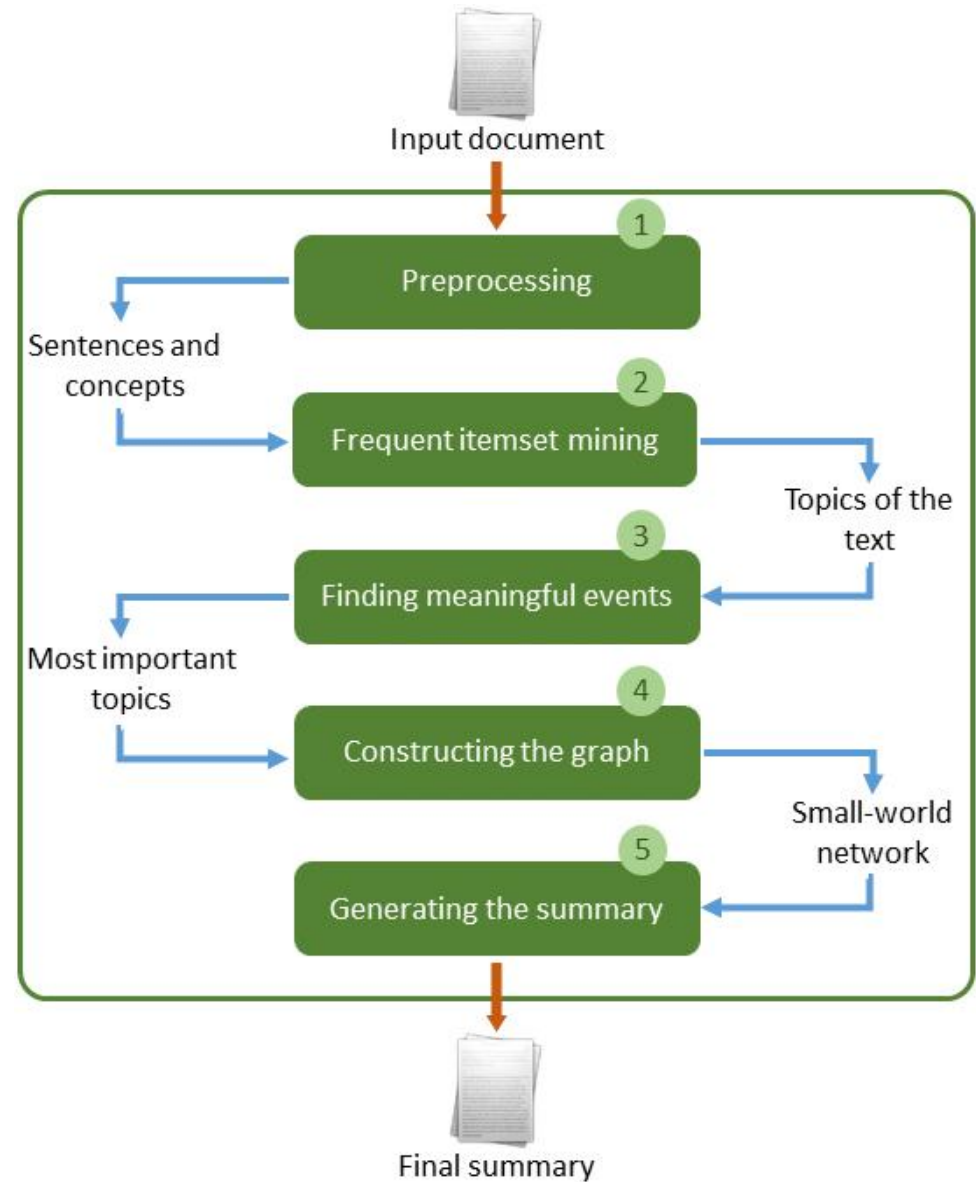
# Kekurangan Chatbot

- *Membunuh Peran Customer Service.* Dengan adanya robot ini pasti akan menggantikan pekerjaan para customer service dan beralih ke Chatbot. Maka dari itu pekerjaan yang awalnya dikerjakan oleh manusia ini pasti akan tersisih dan tergantikan.
- *Tidak bisa Menjawab Keluhan Secara Menyeluruh.* Dengan menggunakan Chatbot, terkadang konsumen mengalami beberapa masalah baru saat menggunakan Chatbot. Masalah tersebut seperti Tidak dapat menjawab semua keluhan. Konsumen tidak akan merasa puas bahkan bisa menambah masalah baru.

# Text Summarization (1)

- Peringkasan teks adalah proses menyaring informasi paling penting dari sebuah teks untuk menghasilkan versi ringkas untuk tugas atau pengguna tertentu.
- Berguna di era ledakan informasi
- Kategori ringkasan teks:
  - Peringkasan satu dokumen / multi-dokumen
  - Peringkasan Ekstraktif / Abstraktif
  - Peringkasan teks yang berfokus pada kueri

# Text Summarization (2)





# Machine Translation

- Menggunakan komputer untuk mengotomatiskan beberapa atau semua proses penerjemahan dari satu bahasa ke bahasa lain.
- Aplikasi terjemahan bahasa seperti Google Translate
- Pengolah Kata seperti Microsoft Word dan Grammarly yang menggunakan NLP untuk memeriksa keakuratan tata bahasa teks.

# Machine Translation (French)


international - Le Monde.fr  
Translated version of http://www.lemonde.fr

Google™ This page was automatically translated from French. View original web page or mouse over text to view original language.

## "It is impossible for journalists to enter Tibetan areas"

Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

**Facts** The Dalai Lama denounces the "hell" imposed since he fled Tibet in 1959  
**Video** Anniversary of the Tibetan rebellion: China on guard  
**Portfolio** | **Reportage** | **Video**



**Accord sur la TVA: "Sarkozy wins case at the worst possible time"**  
The European finance ministers reached on Tuesday to a compromise allowing the reduction of VAT rates in some sectors, including catering.  
**Record** Mixed reactions after the European agreement (reduction in VAT)

**Face aux déficits, la hausse paraît inéluctable**  
Le gouvernement exclut une augmentation des impôts. Philippe Séguin tire la sonnette d'alarme.  
**Infographie** Finances publiques: les gouvernements  
**Les faits** La crise active le débat fiscal  
**Eclairage** | **Compte rendu**

# Machine Translation (Japanese)


asahi.com: 朝日新聞社の連日ニュースサイト  
Translated version of http://www.asahi.com

Google™ This page was automatically translated from Japanese. View original web page or mouse over text to view original language.

## Business

### Latest News

- The exchange of financial stocks fell slightly prominent lower**  
12 stocks in Tokyo, ahead of sell orders from the backlash of higher yesterday, with slightly lower values. Nikkei ... (11:13) [Full article]
- Negotiation and integration of Japan Sompo Japan 興亜 to aggregate in three large camps**  
Sompo Japan Insurance and it's five to start the negotiations for the merger of NIPPONKOA Insurance Co., Ltd. No. 12, 2007, minutes ... (10:33) [Full article]



New Prius

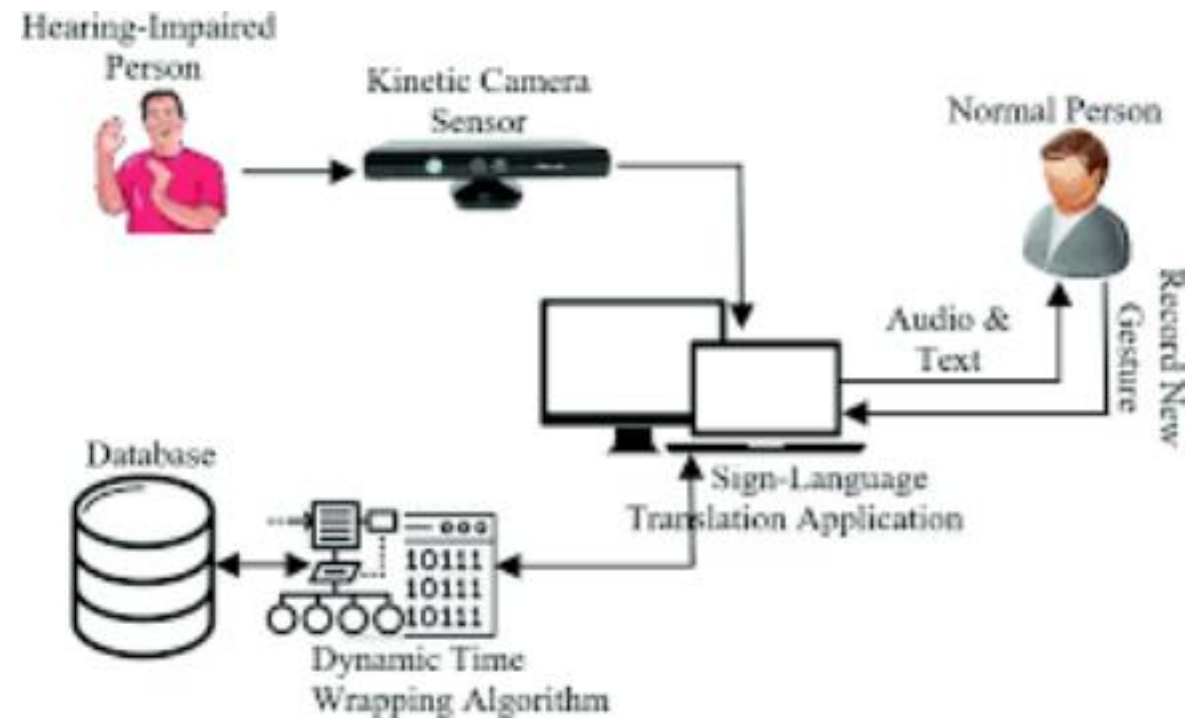
# Aplikasi NLP Lain (1)



NLP juga digunakan untuk filter email. Filter spam telah ada selama beberapa waktu sekarang, tetapi klasifikasi email Gmail adalah salah satu aplikasi NLP yang lebih baru. Berdasarkan konten email yang masuk, Gmail sekarang juga mengenali mana dari tiga kategori (utama, sosial atau promosi) email tersebut berasal. Ini membantu pengguna menentukan email mana yang penting dan membutuhkan respons cepat, dan email mana yang mungkin ingin mereka hapus.

# Aplikasi NLP lain (2)

NLP dapat digunakan untuk menyederhanakan informasi pasien atau untuk aplikasi yang mengubah bahasa isyarat menjadi teks. Ini memungkinkan orang tunarungu untuk berkomunikasi dengan orang yang tidak tahu bagaimana menggunakan bahasa isyarat.



# Aplikasi NLP Lain (3)



Siri berisi :

- Speech recognition
- Language analysis
- Dialog processing
- Text to speech

# Kesimpulan (1)

- Perkembangan teknologi yang sangat pesat diiringi dengan penggunaan internet yang sudah semakin umum, NLP kini telah memiliki peran penting dalam mendukung interaksi antar manusia dan mesin dalam kehidupan sehari-hari.
- Dimulai dari kegiatan sehari-hari hingga tugas-tugas dalam industri tertentu seperti bisnis, kesehatan dan informasi teknologi, NLP dalam tugasnya mempelajari bahasa manusia dengan cara menganalisis sintaksis dan semantis.



# Kesimpulan (2)

- Ada banyak aplikasi NLP di dunia digital dan daftar ini akan tumbuh seiring bisnis dan industri merangkul dan melihat nilainya. Sementara sentuhan manusia penting untuk masalah komunikasi yang lebih rumit, NLP akan meningkatkan kehidupan kita dengan mengelola dan mengotomatiskan tugas-tugas yang lebih kecil terlebih dahulu dan kemudian tugas-tugas yang rumit dengan inovasi teknologi.

# Kesimpulan (3)

- Pemrosesan Bahasa Alami memainkan peran penting dalam mendukung interaksi mesin-manusia.
- Dengan semakin banyaknya penelitian yang dilakukan di bidang ini, kami berharap untuk melihat lebih banyak terobosan yang akan membuat mesin lebih pintar dalam mengenali dan memahami bahasa manusia.





**NLP** Natural  
Language  
Processing

Brand

Reputation

CRM

Quality

**Ai**