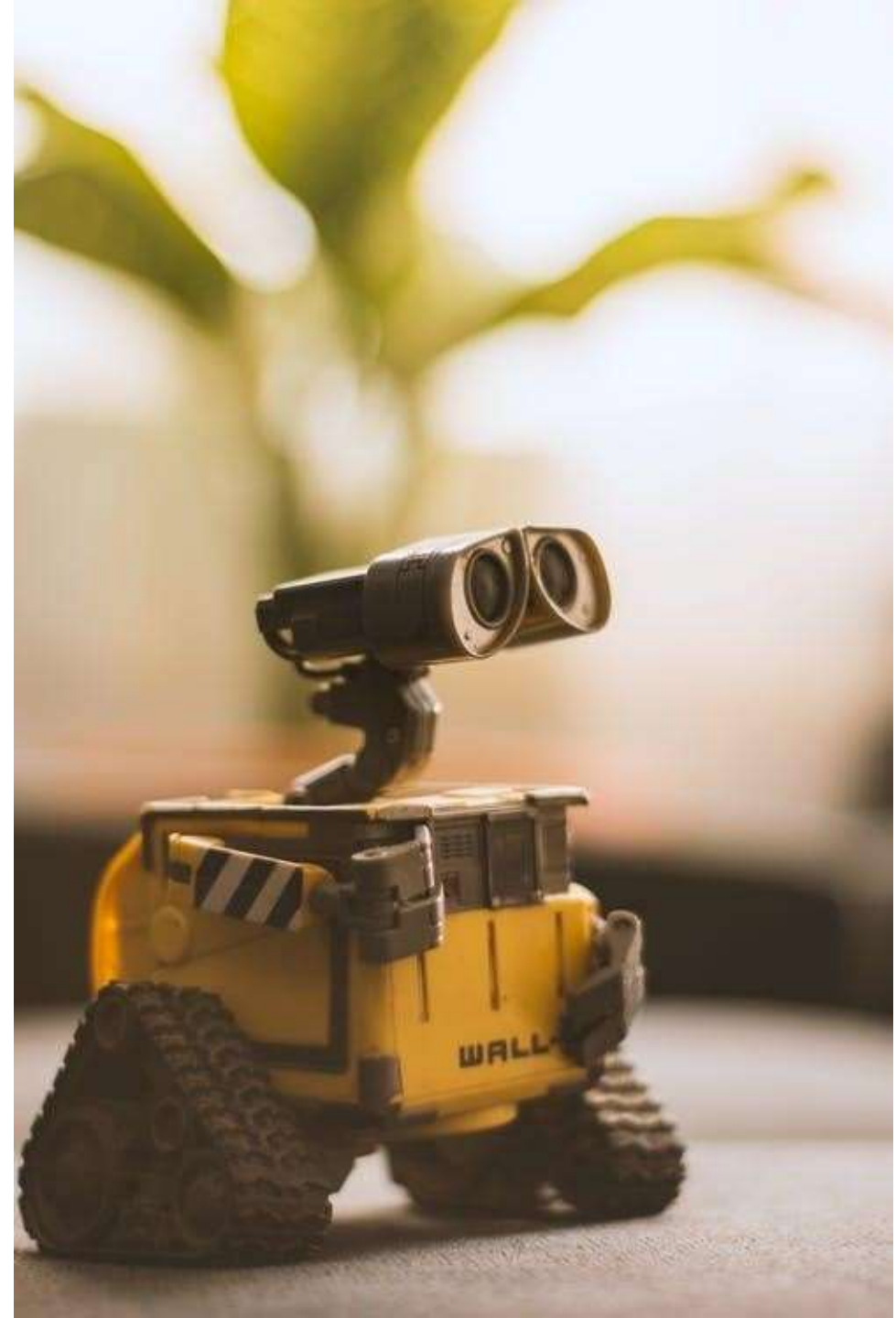




# AI for Text Analytics

Fariz Darari, Ph.D.





***Arek Malang budal nang Berlin taun 2011***

# Fariz Darari, Ph.D.

Assistant Professor at Faculty of CS, UI  
Co-Director at Tokopedia-UI AI Center  
Training Manager at Pusilkom UI

Education:

Libera Università di Bolzano, Italy  
Technische Universität Dresden, Germany  
Universitas Indonesia

Awards:

SWSA Distinguished Dissertation  
Enno Heidebroek Award  
EMCL Best Thesis

**Spree River, Berlin, circa 2011**

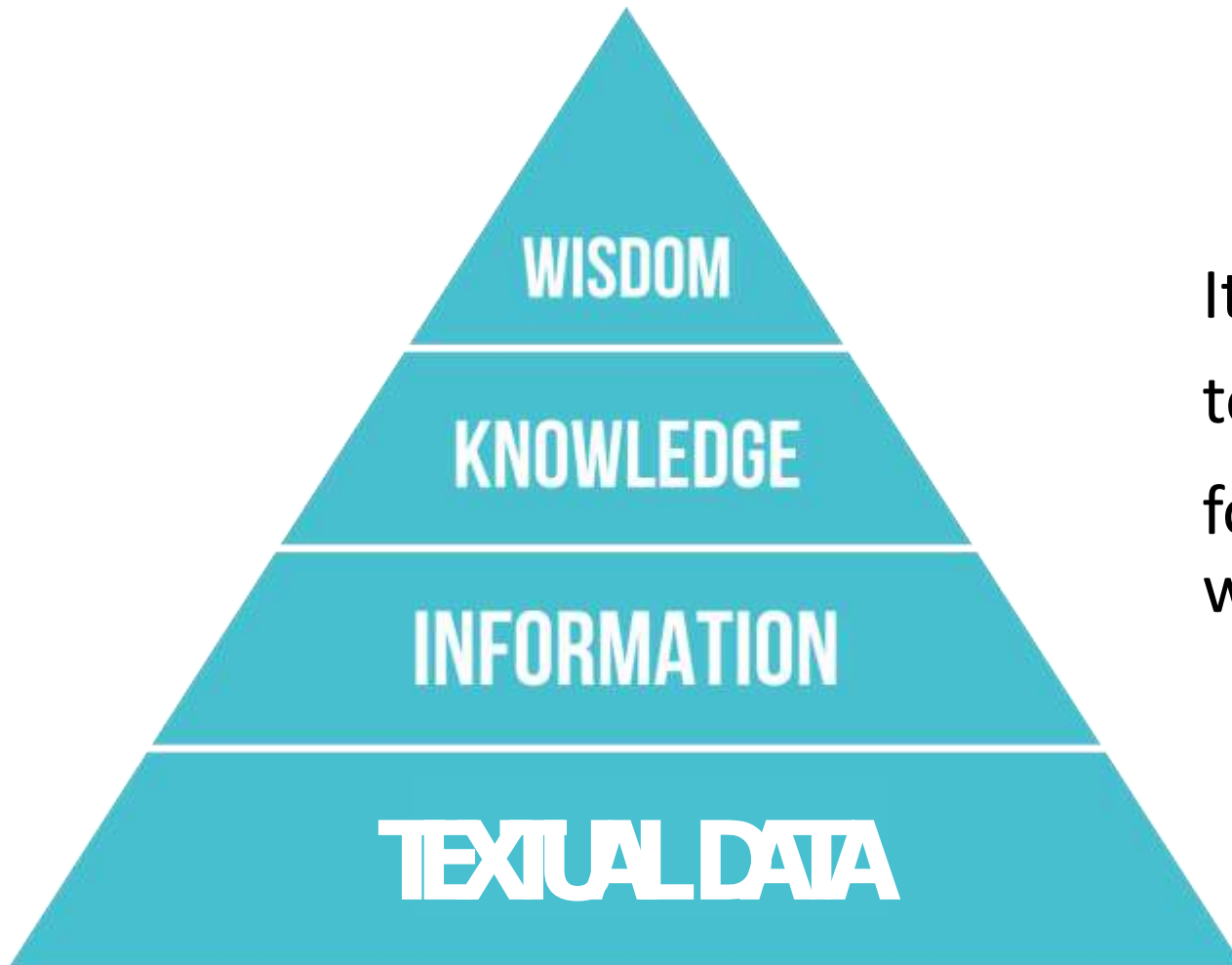
Salam kenal untuk  
rekan-rekan ASN se-Jawa Timur semua! 🤝

Salam hangat dari Universitas Indonesia!



What is text analytics?

# Text Analytics

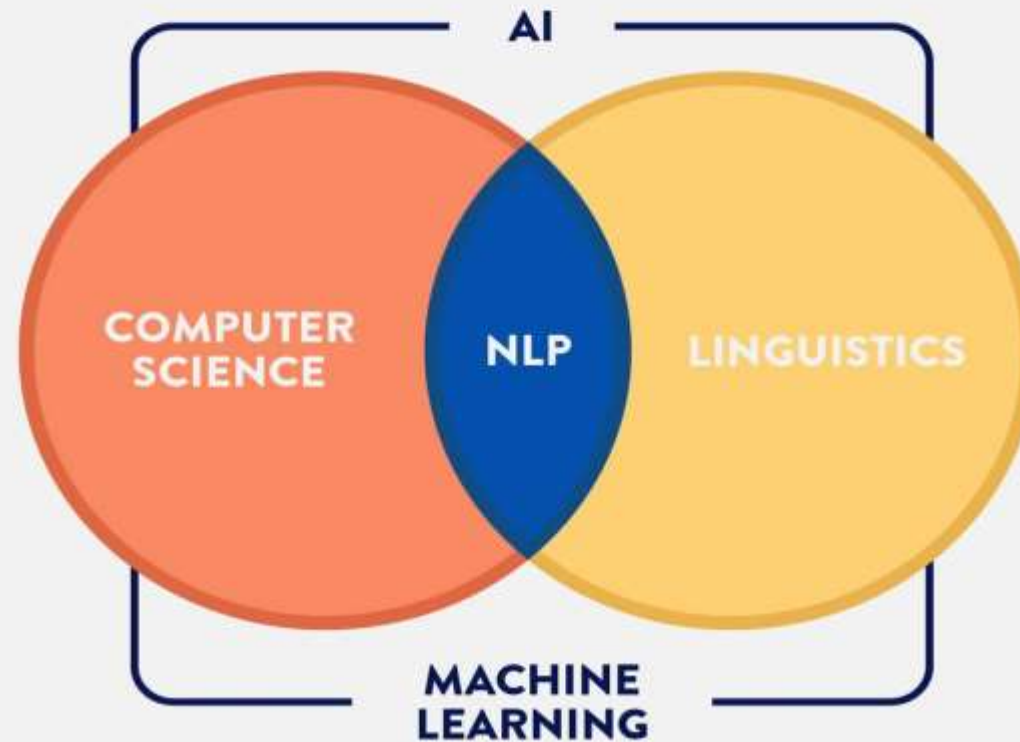


It is the process of using computers, to read and understand texts, for business and organizational wisdom and actionable insights.

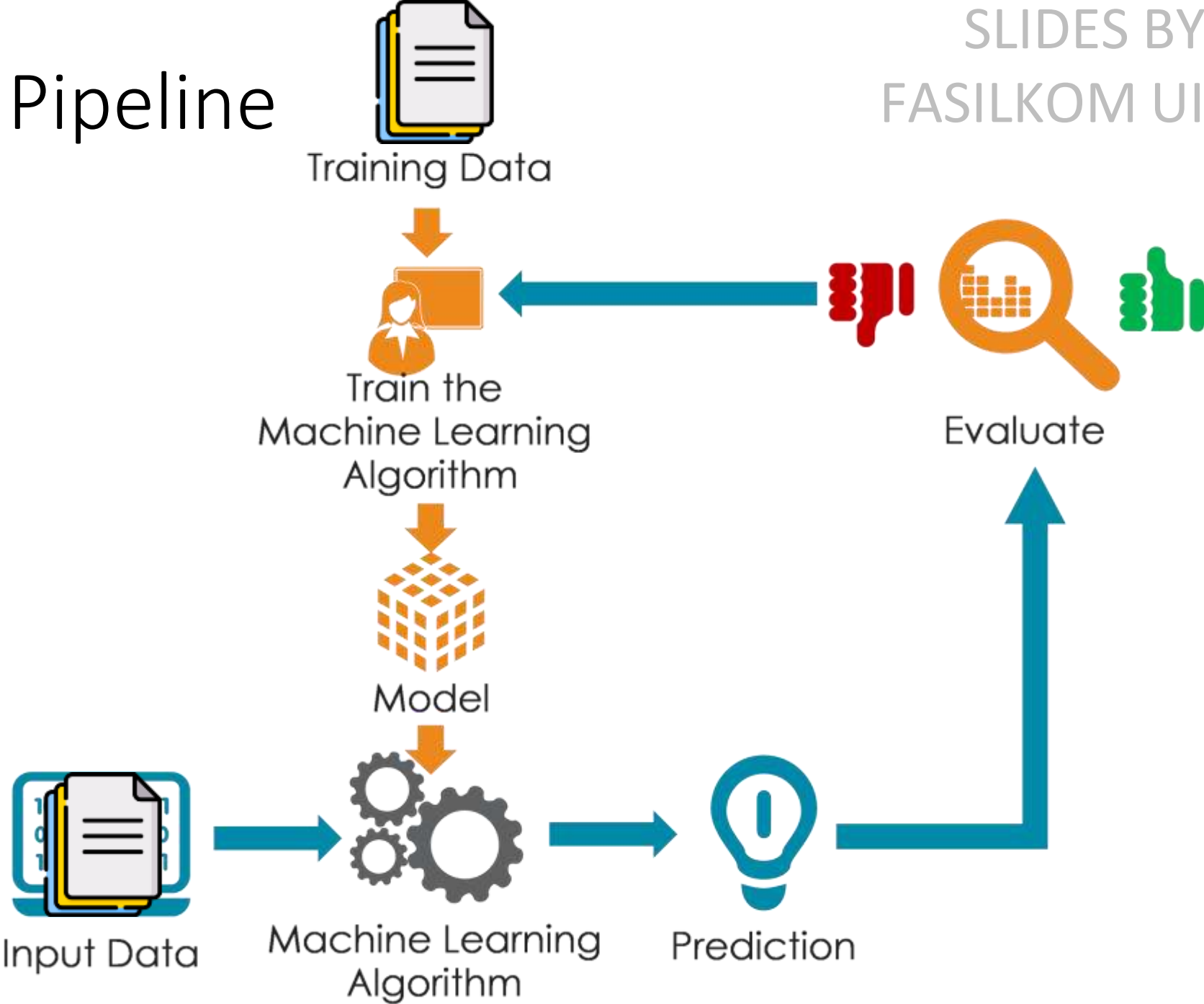
# Text Analytics relies on NLP

## WHAT IS NATURAL LANGUAGE PROCESSING?

— The interdisciplinary field of computer science and linguistics.  
NLP is the ability for computers to understand human language.



# Text Analytics Pipeline



# Text analytics use cases





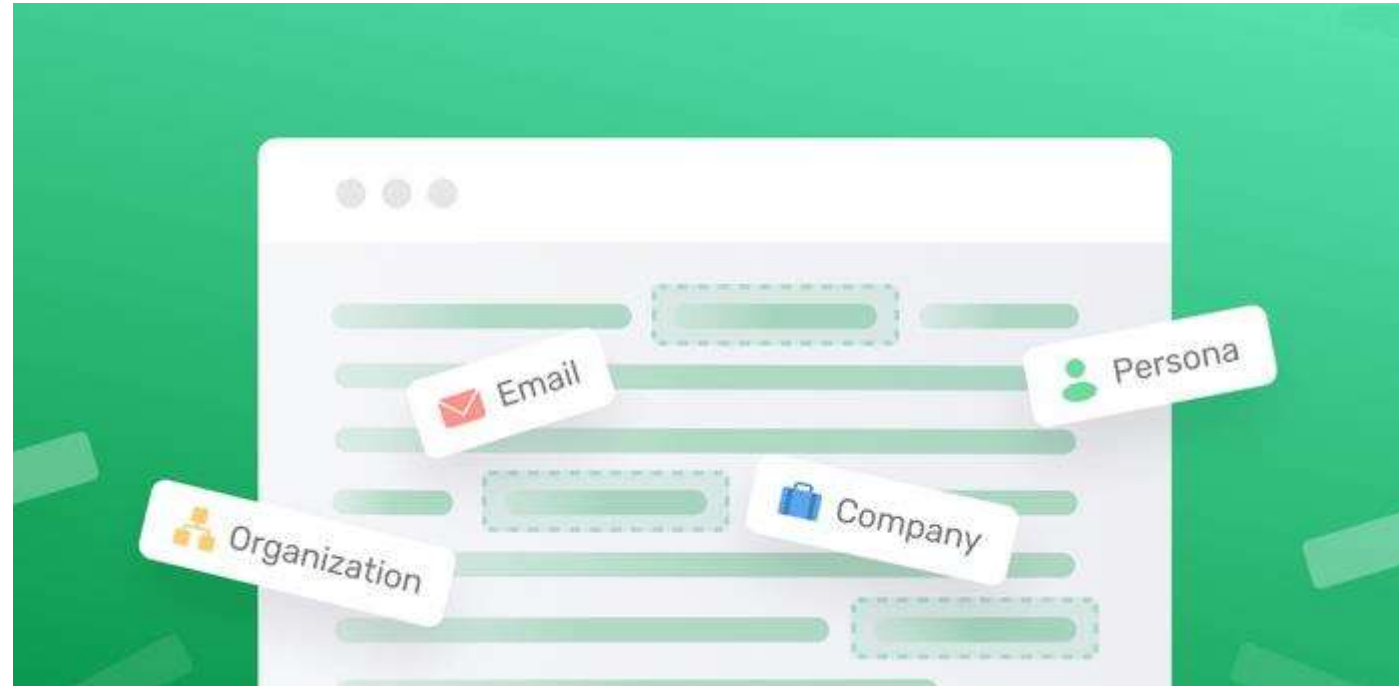
# Word Cloud: DIY

Let's create your own word cloud in just two steps:

- 1) Find a text then copy-paste the text into:  
<https://worditout.com/word-cloud/create>
- 2) Now click: **Generate**

Tips: On the word cloud result, you can click **Regenerate** to change the word cloud style to suit your personality 😎

# Entity Extraction



Entity extraction enables computers to automatically identify or extract entities, like person, organization, event, location, etc

# Entity Extraction: DIY

Let's perform entity extraction easily:

- 1) Find a text then copy-paste the text into:  
<https://www.textrazor.com/demo>
- 2) Now click: **Analyze**
- 3) The extracted entities will appear in bold texts.

# Entity Extraction: DIY

Indonesia is the most beautiful country in the world. I particularly love Surabaya. The people there are so friendly.

Analyze

Clear

[Advanced Options](#)

# Entity Extraction: DIY

Indonesia is the most beautiful country in the world.

Words Phrases Relations Entities Meaning Dependency Parse

---

I particularly love Surabaya.

Words Phrases Relations Entities Meaning Dependency Parse

---

The people there are so friendly.

Words Phrases Relations Entities Meaning Dependency Parse

# Entity Extraction: DIY

I particularly love

**Surabaya**

Words

Phrase

**Normalized Entity Id:** Surabaya

**Normalized English Entity Id:** Surabaya

**Wikipedia Link:**

<http://en.wikipedia.org/wiki/Surabaya>

# Entity Extraction: DIY



WIKIPEDIA  
The Free Encyclopedia

## Surabaya

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

*For other uses, see [Surabaya \(disambiguation\)](#).*

**Surabaya** (Javanese: ꦱꦸꦫꦧꦪ or ꦱꦸꦫꦧꦪꦗꦱꦸꦫꦧꦪ; Javanese pronunciation: [surɔbɔjɔ]; Indonesian pronunciation: [suraˈbaja] listen<sup>ⓘ</sup>) is the capital city of the Indonesian province of East Java and the second-largest city in Indonesia, after Jakarta. Located on the northeastern corner of Java island, on the Madura Strait, it is one of the earliest port cities in Southeast Asia. According to the National Development Planning Agency, Surabaya is one of the four main central cities of Indonesia, alongside Jakarta, Medan, and Makassar.<sup>[8][9]</sup> The city has a population of 2.87 million within its city limits at the 2020 census and 9.9 million in the extended Surabaya metropolitan area, making it the second-largest metropolitan area in Indonesia.<sup>[10]</sup>

The city was settled in the 10th century by the Kingdom of Janggala, one of the two Javanese kingdoms that was formed in 1045 when Airlangga abdicated his throne in favor of his two sons. In the late 15th and 16th centuries, Surabaya grew to be a duchy, a major political and military power as well as a port in eastern Java, probably under the Majapahit empire.<sup>[11]</sup>



# SENTIMENT ANALYSIS



## NEGATIVE

Totally dissatisfied with the service. Worst customer care ever.



## NEUTRAL

Good Job but I will expect a lot more in future.



## POSITIVE

Brilliant effort guys! Loved Your Work.

# Sentiment Analysis: DIY

## Classify Text

The food was great but I hated the waiters.

Classify

👍 **Success** Show REST XML URL

**negative**

53%

**positive**

47%

<https://www.uclassify.com/browse/uclassify/sentiment?input=Text>

# Text Classification

SLIDES BY  
FASILKOM UI



SLID

FASILKOM UI

# Text Classification: DIY

## Classify Text

Cristiano Ronaldo versus Lionel Messi - who is the GOAT? It is the biggest conundrum in football since the 'Diego Maradona or Pele?' debate.

The question of who is better between the Portuguese phenomenon and the empyrean Argentine has dominated the minds of football fans everywhere for the

Classify

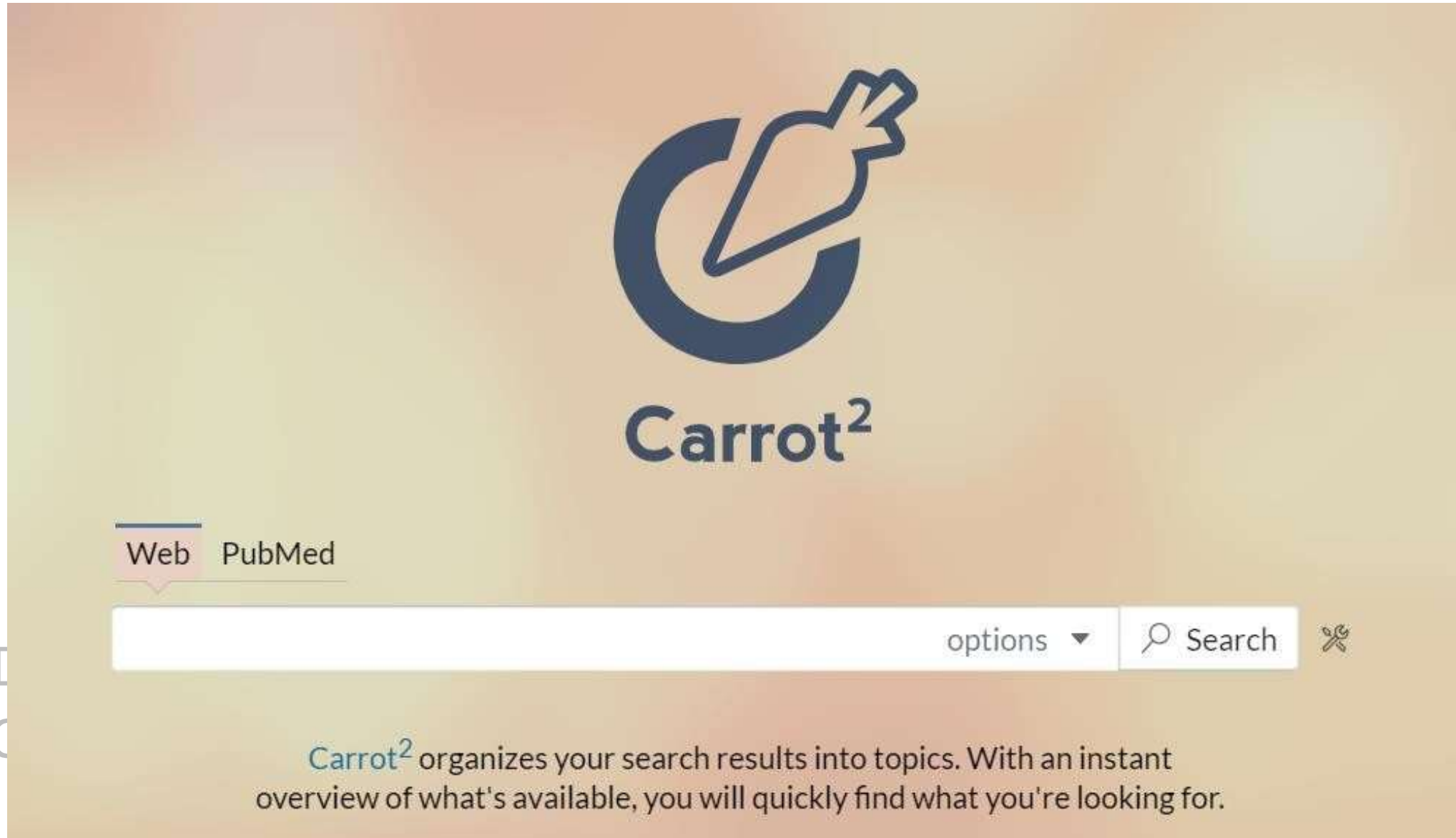
👍 **Success** Show [REST XML URL](#)

**Sports**

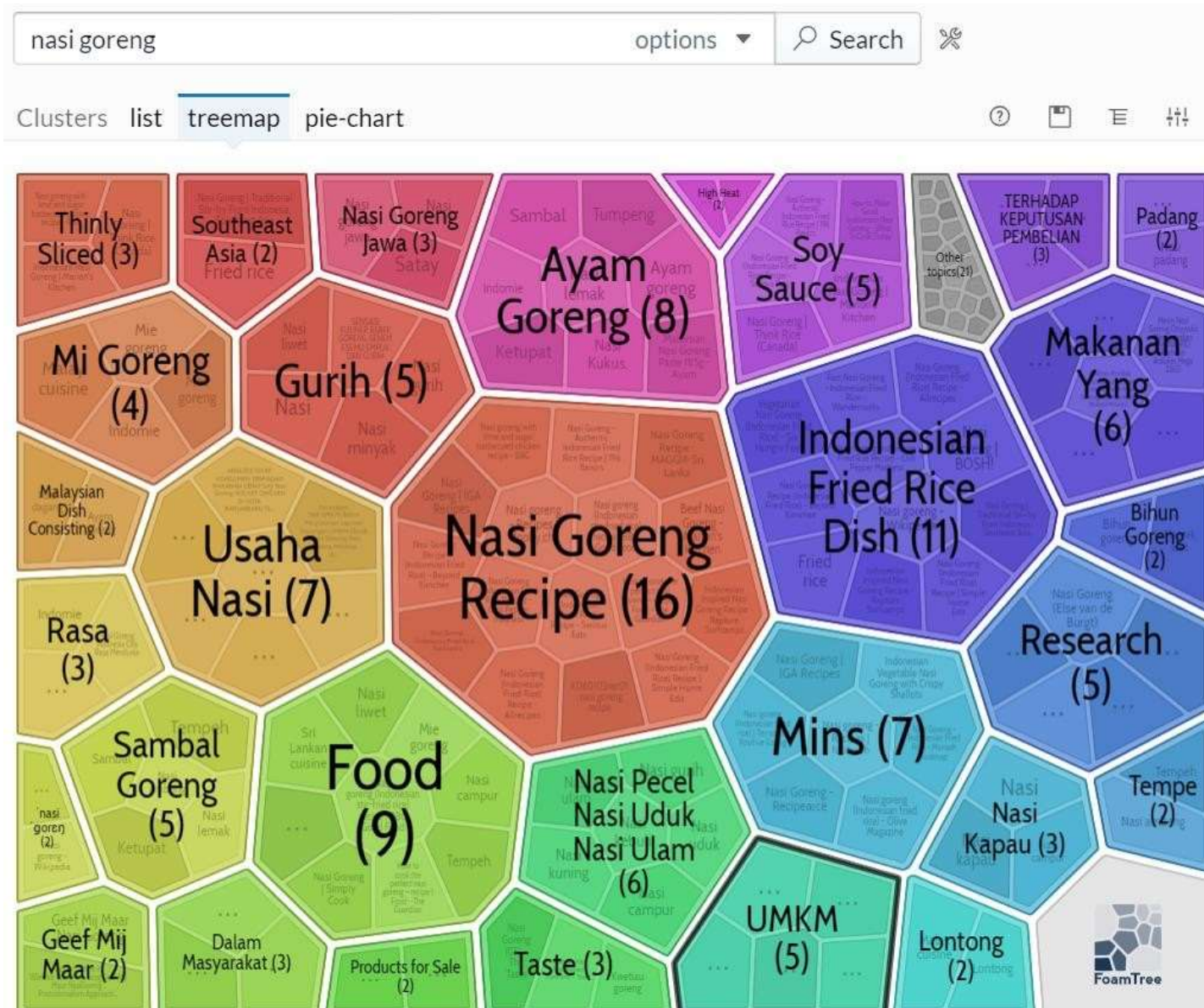
100%

<https://www.uclassify.com/browse/uclassify/topics?input=Text>

# Text Clustering: Reveal Underlying Topics



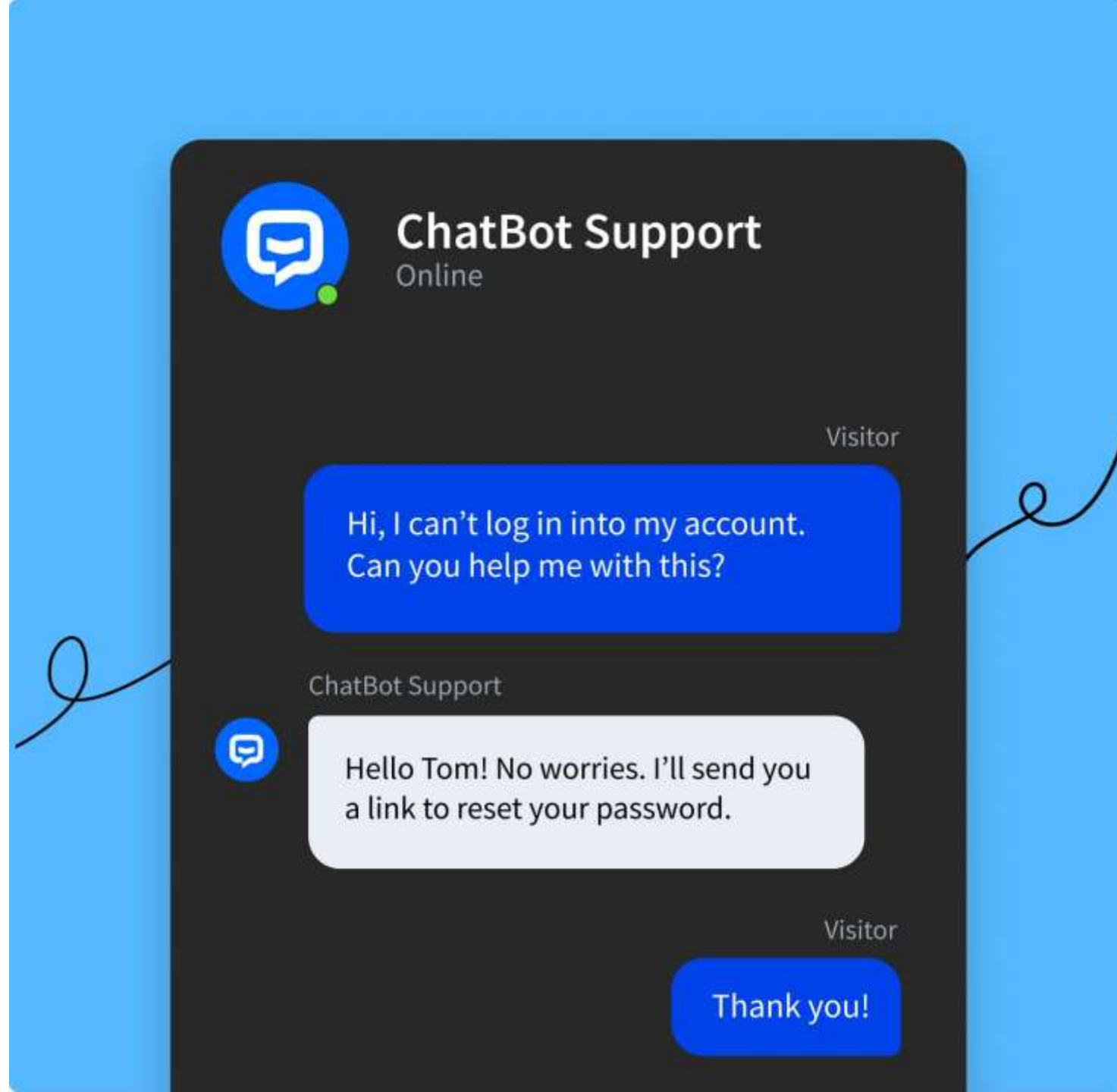
# Text Clustering: Reveal Underlying Topics



## Text Clustering: Reveal Underlying Topics (DIY)

Let's perform text clustering the quick way:

- 1) Go to Carrot at:  
<https://search.carrot2.org/#/search/web>
- 2) Enter your own keywords you want to do clustering, for example, nasi goreng, speak bola, etc
- 3) Choose the tab **treemap** for a nice clustering visualization

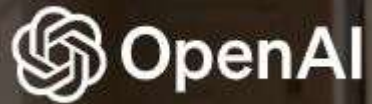




## Chat Bot (Storyteller)



Can you make a bedtime story in Indonesian language about the YouTube channel of Daswer Gaming, which gets more and more viral days by days? The channel is not only about gaming but also cats and creative ways of studying school subjects. The story must include how Daswer, the owner of the channel, is able to make benefits of his channel's popularity for social good (like helping the orphans). Thank you!



# ChatGPT



Can you make a bedtime story in Indonesian language about the YouTube channel of Daswer Gaming, which gets more and more viral days by days? The channel is not only about gaming but also cats and creative ways of studying school subjects. The story must include how Daswer, the owner of the channel, is able to make benefits of his channel's popularity for social good (like helping the orphans). Thank you!



Pada suatu hari di desa kecil yang sunyi, tinggalah seorang pemuda bernama Daswer. Daswer adalah seorang pemuda yang cerdas, penuh imajinasi, dan memiliki cinta yang mendalam terhadap permainan video. Ia juga memiliki seekor kucing lucu bernama Kupu yang selalu menemaninya dalam petualangan hidup.

Daswer suatu hari mendapatkan ide brilian. Ia memutuskan untuk membuat saluran YouTube yang berfokus pada permainan video, kucing, dan cara kreatif dalam belajar pelajaran sekolah. Daswer memberi nama salurannya "Daswer Gaming".

Awalnya, saluran YouTube Daswer Gaming tidak begitu terkenal. Namun, Daswer tidak menyerah. Ia terus berkreasi dan mengunggah video-video menarik yang menampilkan keterampilan bermainnya, kejenakaan Kupu, dan cara unik untuk belajar matematika, sains, bahasa, dan masih banyak lagi.

Lama kelamaan, video-video Daswer mulai menarik perhatian banyak orang. Orang-orang di seluruh dunia mulai mengunjungi salurannya, menyukai, dan berlangganan. Jumlah pengikut dan penontonnya pun meningkat dengan cepat. Daswer Gaming menjadi viral!

SLIDES BY  
FASILKOM UI

See the continuation at:

<https://chat.openai.com/share/0d66449e-125b-4e9e-87cc-f0a4f1908454>

SLIDES BY  
FASILKOM UI

With the additional help of text-to-speech on Google Translate and Veed.io, I can generate this YouTube video of Daswer Gaming Bedtime Story





**Fariz**  
*A regular user of ChatGPT for generating bedtime stories for his son :-)*

**Sam Altman**  
*CEO of OpenAI (Company behind ChatGPT)*



Sam Altman visited Indonesia on June 14, 2023 @ Hotel Indonesia Kempinski Jakarta

## Chat Bot (Social Media Specialist)



Berikut ini adalah informasi pokok dari suatu peristiwa, silakan catat informasinya terlebih dahulu:



Apa: BPSDM Jawa Timur menyelenggarakan Webinar ASN Belajar dengan topik "Peluang dan Tantangan Artificial Intelligence (AI) untuk Mendukung Kinerja ASN"

Siapa: BPSDM Jawa Timur

Narasumber: (1) Nur Aini Rakhmawati, Ph.D (Dosen Teknik Informasi, ITS Surabaya); (2) Dr.rer.nat. I Made Wiryana, S.Si., S.Kom., M.Sc (Dosen Universitas Gunadarma Jakarta, Ahli Kompetensi Keamanan Siber, dan Ketua Indonesia AI Research Consortium); dan (3) Fariz Darari, Ph.D (Dosen di Fakultas Ilmu Komputer UI, peneliti di bidang AI & Knowledge Graph, dan Co-Director Tokopedia-UI AI Center of Excellence).

Kapan: Kamis, 6 Juli 2023, pukul 09.00-12.00

Di Mana: Virtual di YouTube

Mengapa: Perkembangan AI yang pesat di berbagai bidang

## Chat Bot (Social Media Specialist)



Buatkan copywriting untuk Instagram dengan bahasa yang santai dan gaul serta emoji menarik untuk acara webinar di atas.



📍 Hai ASN Jawa Timur! Ada kabar seru nih buat kamu semua! 🌟 BPSDM Jawa Timur bakal menggelar Webinar keren yang nggak boleh kamu lewatkan! 🚀

📅 Catat tanggalnya, ya! Kamis, 6 Juli 2023, pukul 09.00-12.00. 🕒 Di sini, kamu bisa dapetin semua informasi tentang "Peluang dan Tantangan Artificial Intelligence (AI) untuk Mendukung Kinerja ASN"! 🤖

🗉 Narasumbernya juara semua, loh! Kita punya ahli AI yang kompeten abis! Kita bakal dengerin Nur Aini Rakhmawati, Ph.D dari ITS Surabaya, Dr.rer.nat. I Made Wiryana, S.Si., S.Kom., M.Sc dari Universitas Gunadarma Jakarta dan Ketua Indonesia AI Research Consortium, serta Fariz Darari, Ph.D dari Fakultas Ilmu Komputer UI dan peneliti di bidang AI & Knowledge Graph! 😎

🌐 Acaranya virtual, jadi kamu bisa ikutan langsung dari rumah aja! Tinggal buka YouTube dan cari channel BPSDM Jawa Timur. Siap-siap dapetin wawasan baru seputar AI yang lagi naik daun!

Jangan sampe ketinggalan, ya! Siapkan diri kamu dan catat tanggalnya. Sampaikan ke teman-temanmu juga! 📣 Yuk, ramaikan Webinar ini bareng-sama kita! 🤝

#WebinarASNJatim #AluntukASN #BelajarAI #BPSDMJatim #SantaiGaulAI

## Conclusions

We live in the exciting era of AI, natural language processing (NLP), and text analytics.

Can we make the best use of these tech advancements to deliver excellent services to the public with respect to our duties as ASN?

**SURE, WE CAN!**





Thank  
you

