

Business applications of Natural Language Processing



le wagon



MACHINE
LEARNING
TOKYO

gengo®
X

LIONBRIDGE AI

Introduction

- Lionbridge x Gengo
- We provide data collection, annotation and validation services to support our customers build their AI models

LIONBRIDGE AI



Cedric Wagrez

- Responsible for AI services in Japan at Lionbridge
- 17 years in Japan in IT industry

Table of Content

What is NLP

Business case examples

Trends and challenges





What is NLP?

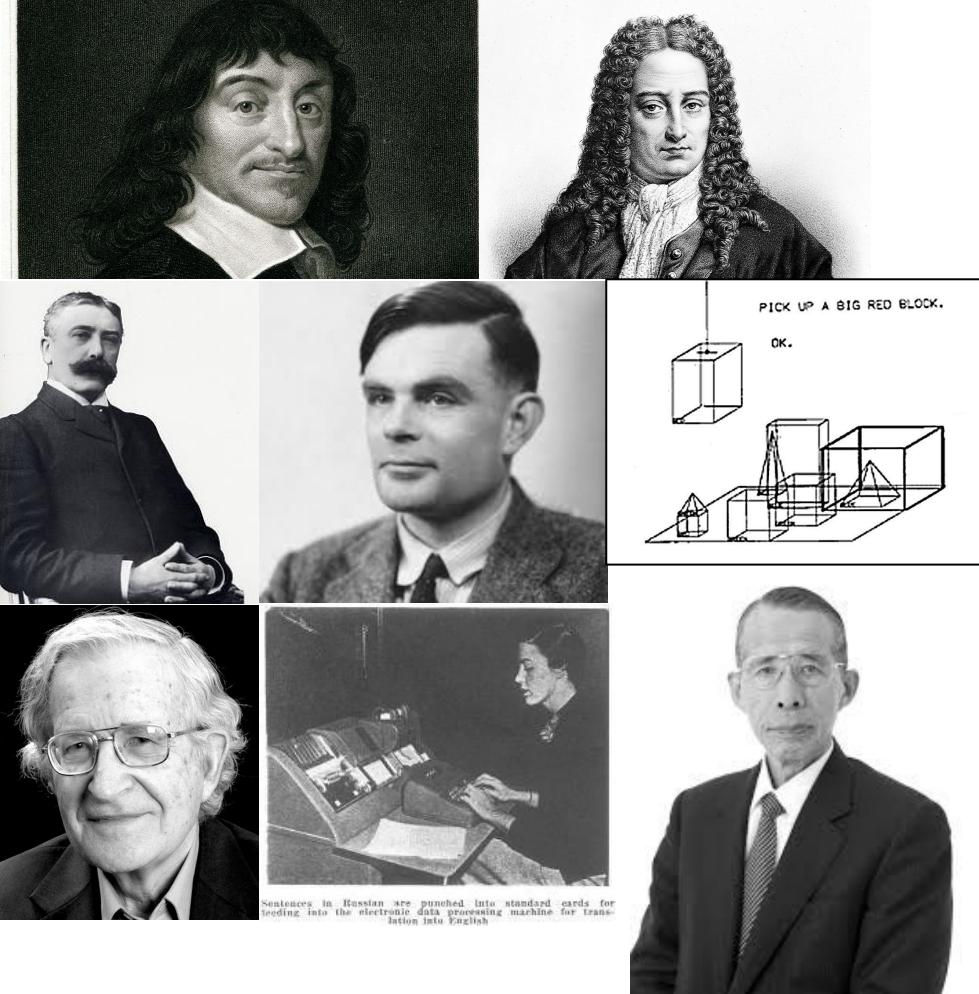
Natural Language Processing [NLP]

Natural Language Processing is a range of technologies that enable machines to understand, analyze and respond appropriately to natural language.

This allows interacting with computers with natural language instead of computer code.

A brief history of.. NLP

- 17th century - Descartes and Leibniz propose independently a way to encode words to allow for mapping between languages
- 1950 Turing test
- 1954 - Georgetown-IBM experiment - first Machine translation ever
- 1957 - Chomsky "Syntactic Structures"
- 1970 - SHRDLU
- 1980s - Japanese companies (NTT, Toshiba, Fujitsu, Mitsubishi, Hitachi, NEC, etc) invest in English <> Japanese machine translation - Nagao Makoto
- 1990s - N-gram, LSTM
- 2010s - acceleration with the breakthrough of deep learning



Accelerated expansion of NLP

2017 June - Transformer Architecture - Google AI

2018 May - Universal Language Model Fine Tuning - fast.ai

2018 November - BERT - Google AI

2019 January - Google's Transformer-XL - Google AI

2019 January - StanfordNLP - Stanford University

2019 February - OpenAI GPT2

2019 June - XLNET - CMUAI

2019 July - PyTorch-Transformers

2019 July - ERNIE - Baidu Research

2019 July - RoBERTa - Facebook Research

2019 August - Spacy-PyTorch-Transformers - SpaCy

2019 August - XML/mBert - Facebook Research

2019 August - MegatronLM - NVidia



Common tasks of NLP

- NER
- Sentiment Analysis
- Text classification
- Information extraction
- Text summarization
- Text generation
- Machine translation
- etc

ORG PERSON LOC

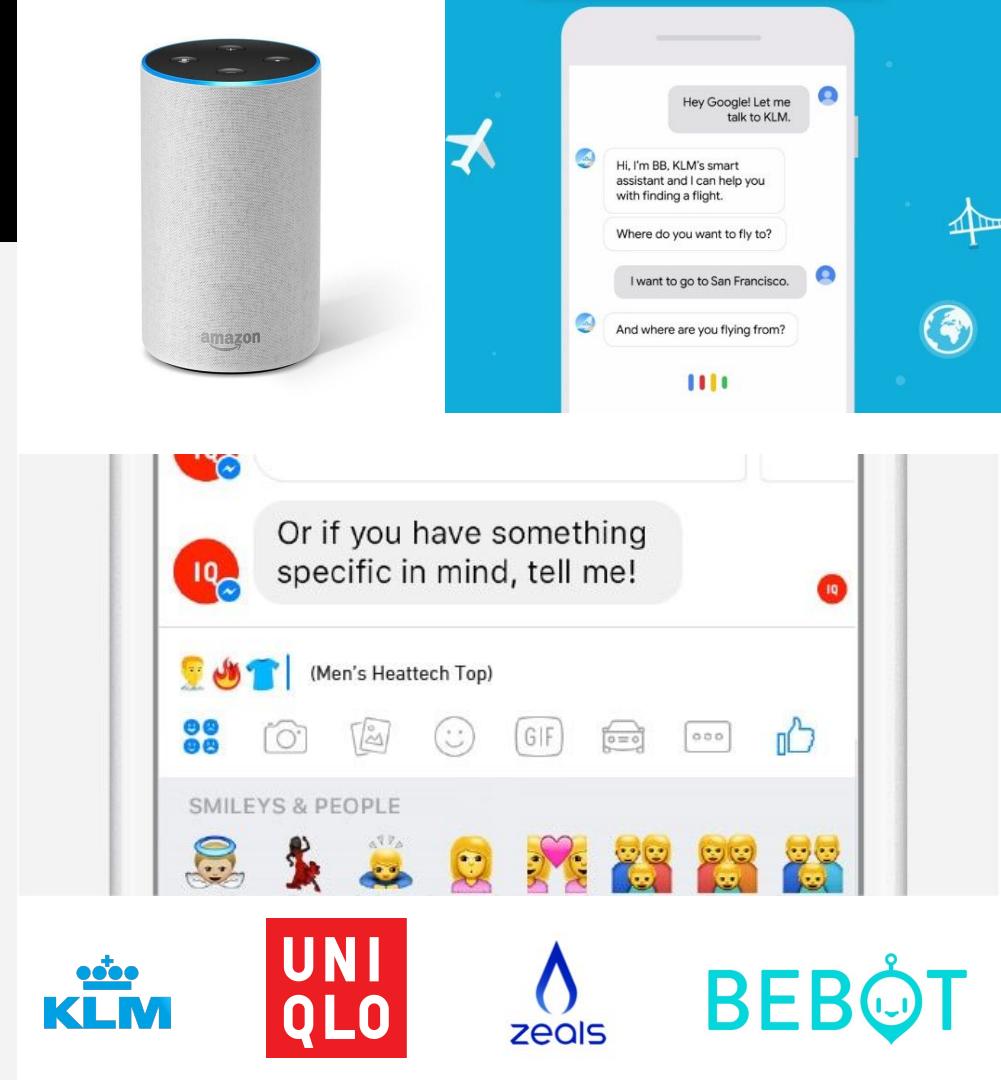
Manchester United produced a stunning comeback from 2-0 down to beat **Newcastle United** - and potentially save manager **José Mourinho** from the sack. Second half strikes from **Juan Mata** and **Anthony Martial** levelled proceedings, before an injury-time winner from **Alex Sánchez** sent the **Old Trafford** crowd into raptures. After the international break, the **Red Devils** travel to **Stamford Bridge** where they face a stern test against unbeaten **Chelsea**.



Applications of NLP

Chatbots

- First thing that comes to mind when thinking of NLP
- It is now easy to create a simple chatbot using Google DialogFlow, Microsoft LUIS and customizing it to your organization needs
- ... but making a great one is a lot of work with conversational design!
- An evolving technology - Google Meena



Support center

- Prioritize contact communications based on topic, sentiment analysis
- Automatic transcription of calls
- Automatically provide relevant articles to the support agent
- Analysis of trends by themes, categories, etc

お客様とオペレーターの会話内容をリアルタイムにテキスト化

会話表示

FAQ検索

Bluetoothのアイコンが表示されないということですね。

検索結果: 100件 スコア順 表示件数

1. Bluetooth機器をパソコンで認識できないため、機器を使用できません。
no_parent-0001-0156
Bluetooth機器を認識できない場合は、パソコンやアプリケーションの設定、環境などを確認します...

2. Bluetoothワイヤレスクロスに対応した機器と、パソコンを接続する手順を教えてください
no_parent-0000-0240
Bluetooth機器とパソコンを接続する方法は、次のとおりです。垂直Bluetooth機器によって、接続の手順が異なることがあります...

FAQ検索結果

FAQ ID : no_parent-0001-0156
質問
Bluetooth機器をパソコンで認識できないため、機器を使用できません
回答
Bluetooth機器を認識できない場合は、パソコンやアプリケーションの設定、環境などを確認します。Bluetooth機器をパソコンで認識できないときは、次の内容を順番にご覧ください。使用環境を確認する省電力機能を解除するBluetooth機器のパソコンへ接続するBluetooth機器を再登録するBluetoothドライバーを再インストールするBluetoothドライバーを更新する~~~~~



<https://webempath.com/>



Machine translation

- The first application of NLP



POCKETALK



<https://www.zoimeet.com/>



ROZETTA



Search engines

- BERT revolution
- Whitepaper about search results improvement with NLP
(Basis Technology - Lionbridge)



2019 brazil traveler to usa need a visa

BEFORE

9:00 google.com

top Washington Post > 2019/03/21

U.S. citizens can travel to Brazil without the red tape of a visa ...

Mar 21, 2019 · Starting on June 17, you can go to Brazil without a visa and ... Australia, Japan and Canada will no longer need a visa to ... washingtonpost.com; © 1996-2019 The Washington Post ...

AFTER

9:00 google.com

USEmbassy.gov > br > Visas

Tourism & Visitor | U.S. Embassy & Consulates in Brazil

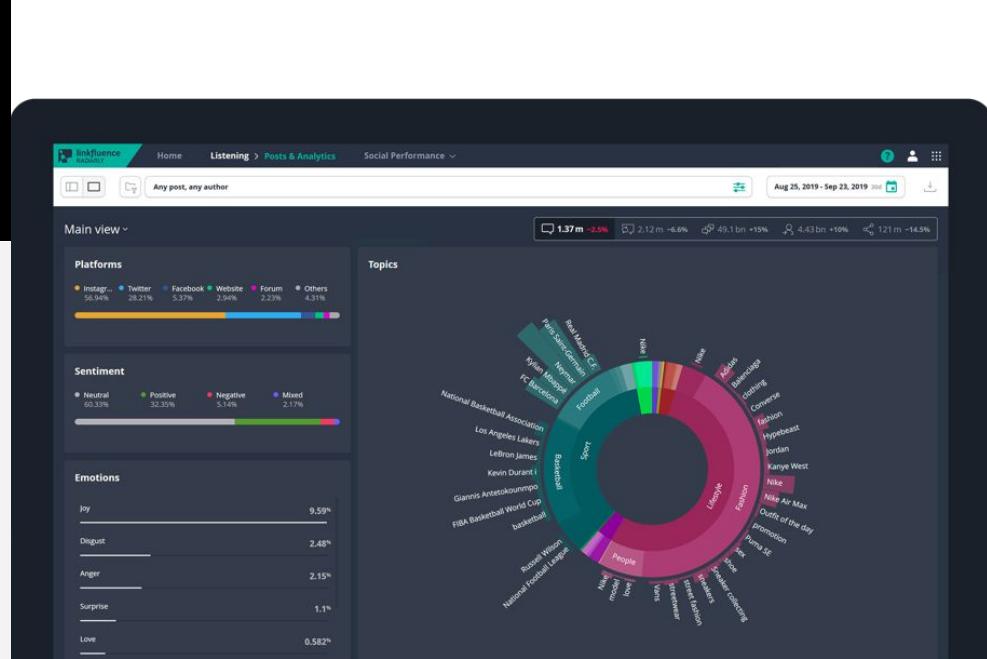
In general, tourists traveling to the United States require valid B-2 visas. That is unless they are eligible to travel visa ...

Google

traveloka

Monitoring of brand and product

- Analyze the reaction towards brands in social media or monitor reviews on EC sites



SIGNAL

Revuze
ANALYST IN A BOX



Where to display an ad

- GDPR and re-marketing
- Display ads based on context (categorization) and also making sure the article is appropriate (not negative, polemic, etc)

Medical assistants

- Transcribing discussions between patients and doctors
- Example of AI being better than physicians sometimes to find risk because of the different ways people with different background may express body feeling and symptoms



Amazon Transcribe Medical

Text generation

- Generating articles, reviews
- Text summarization

AI is Not the Enemy

In many ways, AI is a natural extension of our ability to communicate. This can be seen in two ways. The first is a direct implication of the second: the ability to read the environment. The ability to recognize that another character is in a certain situation is a direct result of having a basic understanding of language.

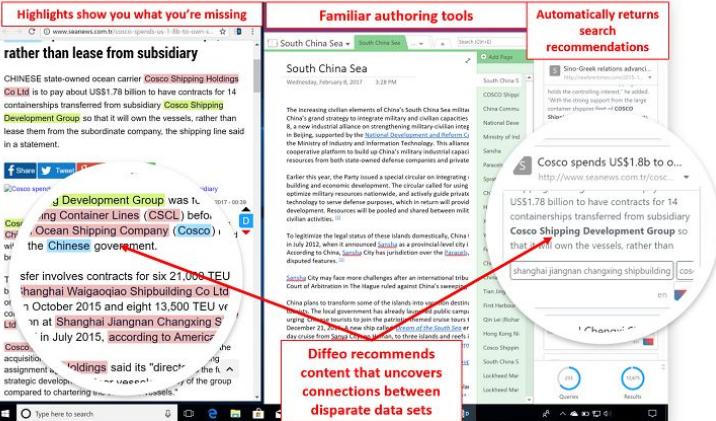
The second is a more abstract concept and one which can be applied to almost any medium: a good writer can tell a good story even if they don't understand how the story works.

But, if we're going to ask whether AI is the enemy, it's important to ask why we are in an environment in which it exists. I think we're on the cusp of something interesting here. For the first time in human history, a new kind of technology has arisen that is creating real value and excitement.



Information extraction

- Extract dynamically information that sleeps in your company systems

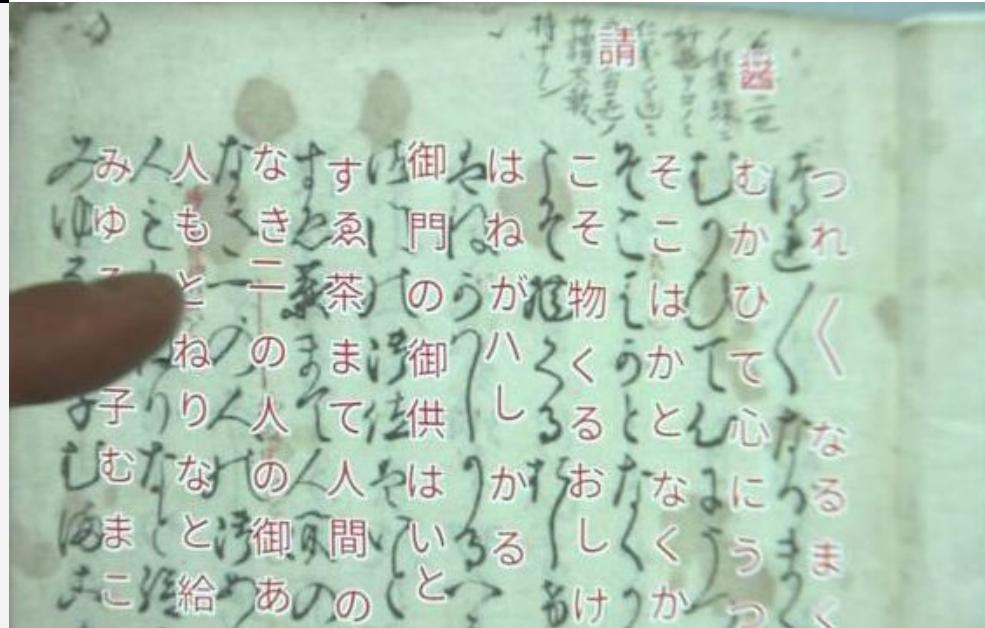


Knowledge Explorer

diffeo

OCR

- OCR is mainly done using computer vision models
- Language Models can be used especially when some characters are illegible or ambiguous. Context in this case may provide suggestions that help identify the most likely character/word when pure computer vision fails



https://www3.nhk.or.jp/news/html/20191202/k1001219856100_0.html

Security

- Immigration checks across languages
- Money laundering prevention



<https://www.basistech.com/>

Resume reviews?

- Scan resumes and find relevant information
- Amazon fail

ideal.



<https://ideal.com/product/screening/>

Mixing it all together

- OCR + Language Model
- Voice command
- Language synthesis



<https://www.orcam.com/en/>

Trends



Technical trends

BERT

Speed of evolution - reading research papers

Size of models now (google Meena has 2.3B parameters).



Global trends

- A wider range of languages
 - Major IT companies are investing in machine translation for languages that were not supported well until now (Africa, Indian dialects)
- Evolution of user experience
 - Voice is starting to take over text
 - Chatbot is the future experience over smartphone screens
- Man-machine interfaces
- Data privacy
 - GDPR
 - Face recognition issues, call center private information, etc



Trends in Japan

- Chatbots
 - Technology and services that have been trending for the last 3 years
- Focus on south-east languages
- An aging population



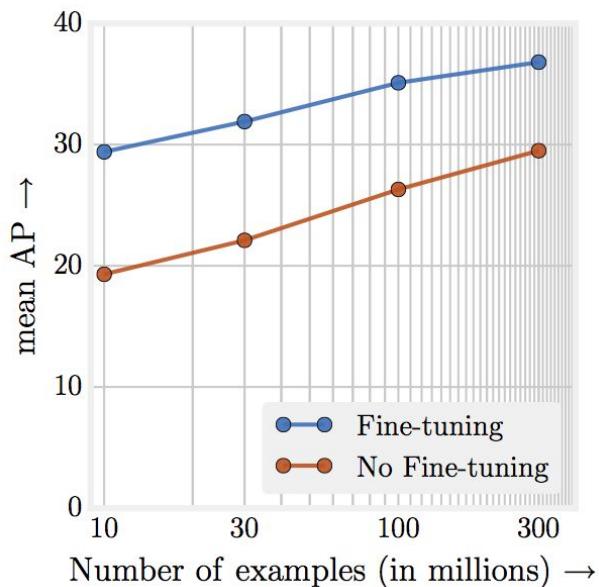
Source: AINow

Data

Multilingual data beyond Chinese and English for text and voice

Quality and volume

Non biased data is very important



Things move fast - keep up to date

- Media!
 - Use twitter
 - Check AI news (media in the US, media in Japan)
- Communities!
 - Join communities and their events! Machine Learning Tokyo, AI Team
 - Japanese - lots of events on Compass, Machine Learning 15 minutes, etc.
- Events
 - AI Expo in April (come say hello)
 - NLP2020
 - Various industry and research event on a regular basis
- Study
 - Read books
 - Take courses - Coursera, fast.ai, join study groups
- Read about use cases
 - e.g. service from ledge AI

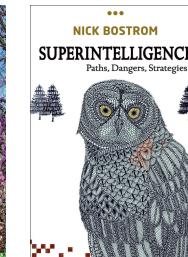
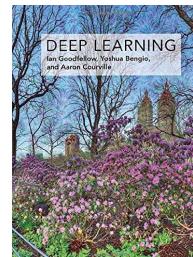
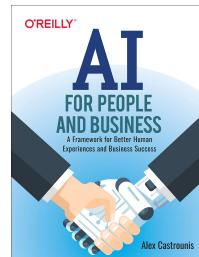


MACHINE
LEARNING
TOKYO



ledge

AINOW



AI Smiley

AI-SCHOLAR.TECH

Towards
Data Science

arXiv.org



deeplearning.ai | coursera

fast.ai

Making neural nets
uncool again

Thank you!



MACHINE
LEARNING
TOKYO



le wagon

LIONBRIDGE AI

cedric.wagrez@gengo.com

@kedbreak136

lionbridge.ai/ja/