

Fact Sheet

ANNEX G

OPEN SOURCE AI LIBRARIES

Artificial Intelligence (AI) has been identified as a frontier technology essential to growing Singapore's Digital Economy. To maintain Singapore's key node status in the global AI ecosystem, IMDA announced in October 2017 that it will spearhead the development of AI areas.

One key focus area is AI libraries, which are collections of Application Programming Interfaces (APIs), source codes, databases and more. IMDA-created libraries can be released as open-source libraries to spur the growth and innovation of AI-related apps in Singapore. Two open-source AI libraries have been created. They are:

- 1) National Speech Corpus V1.0
- 2) Intelligent Sensing Toolbox

National Speech Corpus V1.0

First announced in November 2017, the first version of the National Speech Corpus (NSC) was released on 22 November 2018. It contains 2,000 hours of locally accented audio and corresponding text transcriptions. There are more than 40,000 unique words within the text transcriptions comprising local words such as "Tanjong Pagar", "ice kacang", or "nasi lemak". The data is made available via the Singapore Open Data Licence¹.

Automatic speech recognition engines use multiple corpus collections (collectively called corpora) to accurately train themselves to interpret spoken words and transcribe them. The NSC thus enables global technology providers to provide speech-related applications such as voice assistants, for use here. The NSC will be continually updated.

Interested parties can request for a copy of the corpus at:

www.imda.gov.sg/NationalSpeechCorpus

Intelligent Sensing Toolbox

In Big Data analysis, machine learning-based engines can analyse large datasets from multiple sources and identify patterns to generate insights.

The Intelligent Sensing Toolbox (IST) is a suite of open-source AI tools and technologies developed by IMDA. This sense-making AI algorithm will offer businesses a plug-and-play,

¹ <https://data.gov.sg/open-data-licence>

open-source source code that can be quickly adapted and layered on top of its existing data analytics system to help it make better decisions. The IST is open sourced on the Apache 2.0 License² and more functionality can be added, in collaboration with industry partners.

IMDA had successfully completed a proof-of-concept of the IST through its Intelligent Data Centre programme with the National Supercomputing Centre in 2017. The proof-of-concept demonstrated then that the IST could identify the probable causes of abnormal behaviours of server nodes in the data centre and thus gain new insights.

The toolbox's algorithms are available at <https://gitlab.com/IMDA-DSL/Intelligent-Sensing-Toolbox>

AI as an enabler in the Digital Economy Framework for Action

IMDA, in collaboration with AI Singapore, has announced initiatives in support of our AI strategy to:

1. Drive industry adoption of AI, and promote understanding of the benefits from implementing AI;
2. Build a vibrant ecosystem of AI developers and companies, users and suppliers, and talented researchers and engineers;
3. Identify specific AI leadership areas to grow industry and R&D capabilities; and
4. Provide clarity and review relevant regulations for an AI-conducive business environment.

² <https://www.apache.org/licenses/LICENSE-2.0>

About Infocomm Media Development Authority (IMDA)

The Infocomm Media Development Authority (IMDA) leads Singapore's digital transformation with infocomm media. To do this, IMDA will develop a dynamic digital economy and a cohesive digital society, driven by an exceptional infocomm media (ICM) ecosystem – by developing talent, strengthening business capabilities, and enhancing Singapore's ICM infrastructure. IMDA also regulates the telecommunications and media sectors to safeguard consumer interests while fostering a pro-business environment, and enhances Singapore's data protection regime through the Personal Data Protection Commission. For more news and information, visit www.imda.gov.sg or follow IMDA on Facebook IMDAsg and Twitter @IMDAsg.

For media clarifications, please contact:

NEUBRONNER Eugene (Mr)
Manager, Communications and Marketing, IMDA
DID: (65) 6211 1182
Email: Eugene_Neubronner@imda.gov.sg

KOH Christopher (Mr)
Senior Manager, Communications and Marketing, IMDA
DID: (65) 6211 1709
Email: Christopher_Koh@imda.gov.sg