|  |
| --- |
|  |

|  |
| --- |
| Input Contribution |
| Meeting ID\* | RDM#51 |
| Title:\* | Use case for data augmentation |
| Source:\* | JaeSeung Song, KETI, jssong@sejong.ac.krMinbyeong Lee, Hyundai Motors, minbyeong.lee@hyundai.com  |
| Date:\* | 2021-09-02 |
| Input related to\* | WI-0105 oneM2M System Enhancement to AI capabilitiesTR-0068 V 0.1.0 |
| Intended purpose ofdocument:\* | [x]  Decision[ ]  Discussion[ ]  Information[ ]  Other <specify> |
| Impacted other TS/TR(s) |  |
| Decision requested or recommendation:\* | Agree for inclusion in TR-0068 |
| Template Version: January 2017 (Do not modify) |

**oneM2M Notice**

The document to which this cover statement is attached is submitted to oneM2M. Participation in, or attendance at, any activity of oneM2M, constitutes acceptance of and agreement to be bound by terms of the Working Procedures and the Partnership Agreement, including the Intellectual Property Rights (IPR) Principles Governing oneM2M Work found in Annex 1 of the Partnership Agreement.

# Introduction

This contribution introduces a new use case for data augmentation.

### -----------------------Start of change 1-------------------------------------------

## 7.2 Use case #x – Data Augmentation for AI/ML services

*Editor’s Note: The section introduces a AI/ML use case that uses IoT data.*

### 7.2.1 Description

Image recognition generally use data augmentation methods. Such as; Medical imaging domain utilizes data augmentation to apply transformations on images and create diversity into the datasets. For example, combining pre-trained deep network architectures with multiple augmentation techniques enabled accurate detection of breast cancer from a very small set of histology images with less than 100 images per class [paper: “Deep convolutional neural networks for breast cancer histology image analysis]. The reasons of data augmentation interest in healthcare are

* Small dataset for medical images
* The use of medical image augmentations helped to improve the results of segmentation of hand radiographs and bone age assessment etc.
* Sharing data is not easy due to patient data privacy regulations
* There are only a few patients whose data can be used as training data in the diagnosis of rare diseases

### 7.2.2 Source

### 7.2.3 Actors

* AI/ML management application: an application managing AI/ML image data
* AI-enabled IoT platform: An IoT platform collects image data from various sources

### 7.2.4 Pre-conditions

* The AI-enabled IoT platform holds a set of good quality image data for AI/ML
* The AI-enabled IoT platform provides features to handle requested image augment techniques
* The amount of collected source images to build a model is not enough.

### 7.2.5 Triggers

* AI/ML management application requests to augment data to build AI/ML model

### 7.2.6 Normal Flow

Figure 7.x.6 illusrates the high-level flows of data augmentation for AI/ML use case, which consists of the following steps:

* Step 1: The AI/ML management application sends a request to the AI-enabled IoT platform to augment source images. The request may include the following information:
	+ Source images
	+ Data augmentation techniques to apply
	+ Additional information for a selected data augmentation technique, for example, the number of images to generate after applying the data augmentation technique.
* Step 2: The AI-enabled IoT platform analyses the received request and stores retrieved information internally
* Step 3: The AI-enabled IoT platform applies the selected data augmentation technique and generates a set of augmented images. The IoT platform stores generated images with their own identifiers.
* Step 4: The AI-enabled IoT platform returns the result to the AI/ML management application. The results may include a summary of the requested data augmentation, for example, the number of generated images, links to access such augmented imanges

### 7.2.7 Alternative Flow

None

### 7.2.8 Post-conditions

The AI-enabled IoT platform has data set for the source images and augmented images from the source.

### 7.2.9 High Level Illustration

###

### 7.2.10 Potential Requirements

1. The oneM2M System shall be able to handle data augmentation requests for AI/ML purposes.
2. The oneM2M System shall be able to generate augmented data resources from a given source data and data augmentation technique.

### -----------------------End of change 1-------------------------------------------