|  |
| --- |
|  |

|  |
| --- |
| 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.

R02:

* Added clarification for the relationship between the two introduced applications
* Revised a potential requirements for clarification

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

# 2 References

The following text block applies.

References are either specific (identified by date of publication and/or edition number or version number) or non‑specific. For specific references,only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

## 2.1 Normative references

*As a Technical Report (TR) is entirely informative it shall not list normative references.*

The following referenced documents are necessary for the application of the present document.

Not applicable.

## 2.2 Informative references

Clause 2.2 shall only contain informative references which are cited in the document itself.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] oneM2M Drafting Rules (<http://member.onem2m.org/Static_pages/Others/Rules_Pages/oneM2M-Drafting-Rules-V1_0.doc>)

[i.2] “Large scale labelled video data augmentation for semantic segmentation in driving scenarios”, Ignas Budvytis, Patrick Sauer, Thomas Roddick, Kesar Breen, Roberto Cipolla; Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 230-237

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

### -----------------------Start of change 2-------------------------------------------

## 7.2 Use case #x – Data Augmentation for Autonomous Driving

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

### 7.2.1 Description

AI/ML provides good solutions to various domains that need image classification recognition, such as CCTV and Home security cameras. Similarly, autonomous driving technologies use AI/ML to detect objects around a vehicle. Typically such AI/ML requires a vast amount of datasets to have an accurate result. However, many AI/ML services for autonomous driving still operates on limited size datasets. Even there exist many datasets to use, labelling each image by hand takes around an hour. Therefore, AI/ML services for autonomous driving use techniques for data augmentation to acquire a large number of datasets from limited size datasets.

Applying data augmentation techniques to collected datasets require high computing power and large data storage. The autonomous driving application typically does not have such resources. IoT platforms serve the autonomous driving application can provide a common function for data augmentation. The application can offload their data augmentation tasks (e.g., applying image processing and storing large datasets) to IoT platforms. As many IoT services using AI/ML techniques require data augmentation function, the introduction of data augmentation function to IoT platforms benefits many intelligent IoT services.

Several reasons of data augmentation interest in autonomous driving are

* Limited size available dataset, e.g., road images, various objects
* The use of data augmentations helped to improve the results
* Sharing data is not easy due to data privacy regulations

### 7.2.2 Source

### Large scale labelled video data augmentation for semantic segmentation in driving scenarios [i.2]

### 7.2.3 Actors

This use case contains actors as follows:

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

AI/ML data management application builds a model for autonomous driving, and autonomous driving application uses the developed AI/ML model to make a decision on various driving situations such as when to stop, turn on/off engine.

### 7.2.4 Pre-conditions

* The AI-enabled IoT platform holds a set of good quality image data for autonomous driving
* 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 data management application requests to augment data to build AI/ML model for autonomous driving

### 7.2.6 Normal Flow

Figure 7.2.9-1 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. Then 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 3: 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 images
* Step 4: The autonomous driving application uses augmented dataset to build a model for autonomous driving, object detection, driving control etc. If the AI-enable IoT platform supports the generation of a model for autonomous driving, the application can download and use the model. Otherwise, the autonomous driving application build its own model using the augmented dataset.

### 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



Figure 7.2.9-1 Data augmentation for autonomous driving

### 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.
3. The oneM2M System shall be able to manage data for AI/ML purposes such as model training and augmentation of training dataset.

### -----------------------End of change 2-------------------------------------------