Impact-Aware Manipulation by Dexterous Robot Control and Learning in Dynamic Semi-Structured Logistic Environments

Data Management Plan (2nd update D6.3)

Dissemination level	Public (PU)
Work package	WP6: Management
Deliverable number	D6.6
Version (D = Draft, F = Final)	F – 1.0
Submission date	17/05/2023
Due date	30/04/2023

www.i.am-project.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871899.

Authors

Authors in alphabetical order		
Name	**Organisation**	**Email**
Maarten JONGENEEL	TU/e	m.j.jongeneel@tue.nl
Jos DEN OUDEN	TU/e	j.h.v.d.ouden@tue.nl

Control sheet

Version history			
Version	**Date**	**Modified by**	**Summary of changes**
0.1	07/04/2023	Jos DEN OUDEN	ToC & first contents
0.11	14/04/2023	Jos DEN OUDEN	New layout & updated all sections
0.14	14/04/2023	Jos DEN OUDEN, Maarten JONGENEEL	Improvement first draft (modification details)
0.5	08/05/2023	Maarten JONGENEEL, Jos DEN OUDEN	Pre-final version ready for peer-review
0.9	17/05/2023	Alessandro SACCON	Peer-review comments addressed
1.0	17/05/2023	Jos DEN OUDEN	Revised version ready for submission, quality check

Peer reviewers		
	Reviewer name	**Date**
Reviewer 1	Alessandro SACCON	17/05/2023

Legal disclaimer

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviation	Definition
DMP	Data Management Plan
DOI	Digital Object Identifier
EC	European Commission
HDF	Hierarchical Data Format
PU	Public
SEO	Search Engine Optimization
WP	Work Package

EXECUTIVE SUMMARY

This deliverable D6.6 Data Management Plan (2nd update D6.3), is an update of the second version Data Management Plan D6.5 and provides some updates regarding plan of the dataset generated in WP1 (I.Model), that is to be used by the entire consortium. Throughout the deliverable the updates are clearly shown, for the reader. Since the dataset itself was already described and published in D1.1 and succeeding D1.4, previous deliverables D6.3 and D6.5 focused on HOW we will make dataset and repo available according to FAIR policy. D1.4 focuses on WHAT we make available (in-depth dataset).

Since the I.AM. project focuses within WP1 on the generation of the I.AM. dataset, we will describe how we intend to share it and the possible re-use within and outside the consortium. Further detailed technical implementation of the dataset was described in D1.1 in M9 (September 2020) and D1.4 in M36 (December 2022), where the contents of the dataset, the location and the first description of how we ensure open access to public were explained.

Since there are no longer any changes in how we provide the data to the public, this deliverable mainly updates with respect to D1.4 the datasets that are added in the time between D1.4 and this deliverable.

1. INTRODUCTION

1.1. I.AM. project background

Europe is leading the market of torque-controlled robots. These robots can withstand physical interaction with the environment, including impacts, while providing accurate sensing and actuation capabilities. I.AM leverages this technology and strengthens European leadership by endowing robots to exploit intentional impacts for manipulation. I.AM focuses on impact aware manipulation in logistics, a new area of application for robotics which will grow exponentially in the coming years, due to socio-economical drivers such as booming of e-commerce and scarcity of labour.
I.AM relies on four scientific and technological research lines that will lead to breakthroughs in modeling, sensing, learning and control of fast impacts:

1. I.Model offers experimentally validated accurate impact models, embedded in a highly realistic simulator to predict post-impact robot states based on pre-impact conditions;
2. I.Learn provides advances in planning and learning for generating desired control parameters based on models of uncertainties inherent to impacts;
3. I.Sense develops an impact-aware sensing technology to robustly assess velocity, force, and robot contact state in proximity of impact times, allowing to distinguish between expected and unexpected events;
4. I.Control generates a framework that, in conjunction with the realistic models, advanced planning, and sensing components, allows for robust execution of dynamic manipulation tasks.

This integrated paradigm, I.AM, brings robots to an unprecedented level of manipulation abilities. By incorporating this innovative technology in existing robots, I.AM enables shorter cycle time (10%) for applications requiring dynamic manipulation in logistics. I.AM will speed up the take-up and deployment in this domain by validating its progress in three realistic scenarios: a bin-to-belt application demonstrating object tossing, a bin-to-bin application object fast boxing, and a case depalletizing scenario demonstrating object grabbing.

1.2. Data management plan background

The I.AM. consortium agreed to join the H2020 Open Research Data Pilot (ORDP), since the consortium had the intention to publish a publicly available impact aware manipulation database.

According to the I.AM. Grant Agreement article 29.3, the ORDP states that the consortium must comply with the following:

"Regarding the digital research data generated in the action ('data'), the beneficiaries must:

1) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:
    (a) the data, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible;
    (b) other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan'.
2) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.
As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data under Point (a)(i) and (iii) if the achievement of the action's main objective would be jeopardised by making those specific parts of the research data openly accessible.
In this case, the data management plan must contain the reasons for not giving access."

In line with these ORDP guidelines, a first Data Management plan (D6.3) was produced by M6 (June 2020), showing how I.AM. would intend to comply with FAIR policy.
The first version of the I.AM. database was thereafter published in M9 (September 2020) as part of the effort of WP1 and made available to public on the delivery of D1.1 "Publication of I.AM. Database". Details on publication and availability can be found in that specific deliverable. Since the I.AM. consortium plans to grow this database during the project, it is a work in progress and will be continuously updated throughout the entire project till at least M36 (December 2022).

The Data Management Plan is currently finally updated again for a last time at M40 (April 2023) now, publishing latest developments in how the consortium makes the data FAIR (Findable, Accessible, Interoperable, Re-usable). Although from the project point of view, this deliverable will show the latest version, the database is still to be updated in the coming months of the project and is intended to be updated as well after finalizing the project, since the database is being open sourced.

1.3. Purpose of the deliverable

This deliverable D6.6 "Data Management Plan (2nd update D6.3)'' provides the final update over the original data management plan (D6.3) that was published in June 2020 (M6) and which was already updated by D6.5 in January 2021 (M30), containing an 2nd update on the implementation of the database generated in WP1, which is being set up and used by the entire consortium and being made available to public now.

Since the data management plan was originally requested due to the publication of the I.AM. database and this database also already was to be described in detail in D1.1 and D1.4 (both also public available deliverables), we build further upon those two deliverables here as well and present here the current final version.
For compliance with FAIR policy, and choices for data server and its first implementation, we refer to earlier deliverables D6.3 and D6.5 in which this is described.
The first detailed technical implementation of the database was already described in D1.1 in M9 (September 2020), and was updated in D1.4 in M36 (December 2022), which explains in more detail the structure and the contents of the database, its contact transition taxonomy, first description of its (online) location, and the first description of how we ensure open access to public and status of the database per December 2022.
As described, the database itself is a continuing growing database, meant to be extended by others outside of the consortium as well, so this deliverable will only showcase the snapshot as of April 2023.

This deliverable provides an update of D6.3, D1.1, D6.5 and D1.4 focusing on the current implementation of the database and mainly updating on some additional databases that are being added since publication of D1.4 in December 2022.

1.4. Intended audience

The dissemination level of D6.6 is 'public' (PU) – meant for members of the Consortium (including Commission Services) and the public. This document is intended to serve as an internal guideline for the entire I.AM. Consortium and provide the consortium's implementation final implementation regarding data management, with a specific focus on the I.AM. Database that is being made openly available.

2. ACCESSING THE DATABASE

With respect to D1.4, no major changes have taken place on how to access the database as an external party. Therefore, we refer the reader to D1.4 (published December 2022 on the I.AM. website here:) for further reading and/or visit the webpage of de database itself for further reading here: https://impact-aware-robotics-database.tue.nl/.

3. CONTRIBUTING TO THE DATABASE

With respect to contributing to the database, some changes have occurred with respect to D1.4. Therefore, in this section, we will repeat the description of the developed software architecture for collecting and storing experimental data according to the structure discussed in Deliverable D1.4 already, so it is clear how data can be contributed to the database after project ends and explain the additions later.

Firstly, the database consists of multiple datasets. The structure of a dataset is also shown in Figure 1. Within this software architecture, we distinguish between data collection (e.g., controlling the experimental setup and collecting raw sensor data) and data storage (e.g., post-processing raw sensor data to enhance human and machine interpretability, providing metadata, and storing the data in a structured way). In the following subsections, we will detail our approach on data collection and storage. We share the software online, which provides an example and possibility for other researchers that want to contribute to the Impact-Aware Robotics Database.

3.1. System Overview

Figure 2 provides an overview of the developed software architecture. A host PC controls the experimental setup and logs the incoming data from the sensors via the data collection application, which also stores the data of each experiment in separate recording folders. Next, via the data storage application, the raw sensor data is processed and stored together with the associated metadata in datasets. These datasets are uploaded to 4TU.ResearchData [4] OpenDAP servers as part of the Impact-Aware Robotics Archive Collection [3]. The custom web application, as discussed in Section , extracts information from the 4TU servers, which is then displayed on the front-end web interface.

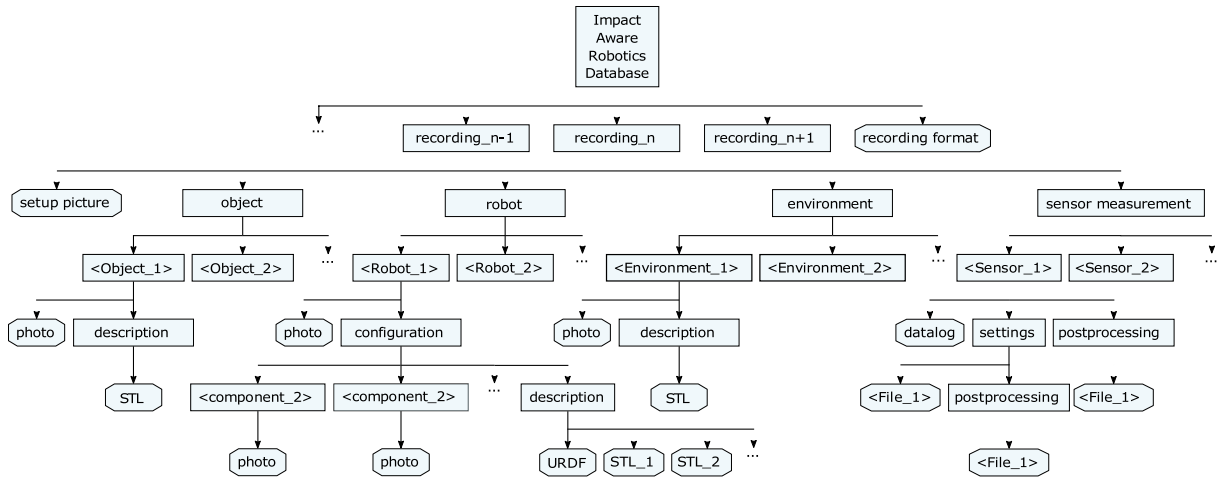

Figure 1: Structure of a dataset. Each dataset contains a set of recordings, as well as a description of the recording format. Each recording is subdivided into object, robot, environment, and sensor measurement subfolders and contains a picture of the setup used during the data recording. This figure is an updated version of Figure 2 from Deliverable D1.1.

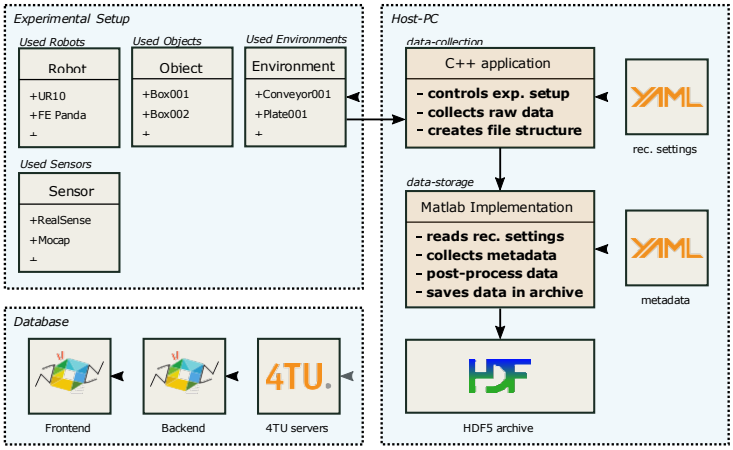

Figure 2: Overview of the software architecture. The host PC contains the data-collection application that takes the recording settings from a configuration YAML file (top right corner) to control the experimental setup, collects the data from the recording, and stores the data in a session/recording file structure. The data-storage application takes this raw data and stores it, together with the associated metadata imported from an external Git-repository (middle right), to an HFD5 dataset file. This file is then uploaded to the 4TU servers from which we display the datasets on the front-end.

3.2. Data Collection

Within the I.AM. project, data is collected at different experimental setups. At the TU/e, two different setups are used for the BOX, TOSS, and GRAB scenarios. At Franka Emika, a setup is used to test intentional impacts (details in deliverable D5.2) related to the GRAB scenario. At other partners, secondary local setups for the GRAB scenario are created. Because experimental data is collected at all these setups, it is important that the software to collect this information is modular, such that the collected data always has the same structure, even though it is recorded on different setups. Furthermore, it is important to document all metadata about the setup, the robot (with its configuration), the objects, and the environment used during the recording of an experiment. Important information to record is also the geometric (dimension, shape) and dynamic (mass, inertia, material) properties of the robot(s), object(s), and environment involved in the experiments, to allow for the reproduction of the experimental conditions. For commercial products (robots, end-effectors, consumer objects), it is requested to describe the exact model/type of product.

With respect to D1.4, now the procedure of data storage has also been made available on a public Gitlab, accessible via: https://gitlab.tue.nl/robotics-lab-public/impact-aware-robotics-database/-/tree/master/data-collection

README.md

The Impact-Aware Robotics Database - Data Collection

Authors
[Maarten Jongeneel, Sander Dingemans, Alessandro Saccon]

The Impact-Aware Robotics Database
[Front-end webpage]

Main webpage of the I-AM project
[i-am webpage]

Introduction

This repository will be used to host the scripts needed to log data of arbitrary sensors at any setup.

Installation

Prerequisites for all platforms

- CMake 2.8.0 or later
- librealsense library
- YAML C++ parcer

If `librealsense` is not installed on your system, please follow these steps to do so:

```
git clone https://github.com/IntelRealSense/librealsense.git
mkdir build
cd build
cmake ..
sudo make clean
sudo make install
```

This should install the librealsense library on your system.

If `yaml-cpp` parcer is not installed on your system, please follow these steps to do so:

```
git clone https://github.com/jbeder/yaml-cpp
cd ./yaml-cpp/
mkdir build
cd build
cmake ..
sudo make install
```

This should install the `yaml-cpp` library on your system.

Linux installation

In order to install this application, clone this git to a desired repository on your system. Then, open a terminal in that repository and execute the following commands:

```
mkdir build
cd build
```

Figure 3: Screenshot of the public Gitlab README file with the instruction on how to use the data collection of the I.AM. dataset.

3.3. Data Storage

The purpose of the data storage application is to post-process and store the recorded (raw) data files together with the associated metadata in a single dataset according to the structure shown in Figure 4. From the recording configuration YAML files is known for each recording what objects, environments, robots, and sensors are used. This information is used to retrieve the associated metadata (e.g., pictures, STL files, robot configuration files) from an external Git repository (middle right of Figure 2) and store this information in the dataset together with the raw and post-processed data. Storing the metadata together with the sensor data ensures the datasets are stand alone and the data can be interpreted by both humans and machines. Once a dataset is created in the form of a HDF5 file, it can be added to the Impact-Aware Robotics Database, for which we detail the procedure in the next section.

With respect to D1.4, now the procedure of data storage has also been made available on a public Gitlab, accessible via: https://gitlab.tue.nl/robotics-lab-public/impact-aware-robotics-database/-/tree/master/data-storage

README.md

The Impact-Aware Robotics Database - Data Storage

Authors
[Maarten Jongeneel, Sander Dingemans, Alessandro Saccon]

The Impact-Aware Robotics Database
[Front-end webpage]

Main webpage of the I-AM project
[i-am webpage]

Introduction

The content of this repository is used to create the archives as presented in the Impact Aware Robotics Database. The raw sensor-data is collected using the data-collection git repository and post-processed using the content of this git repo to be converted into an archive. The main purpose of this repository therefore is to post-process the raw-data into user-interpretable format and add collect the metadata to be stored into the archives.

The archives are stored in HDF5 format (see the HDF-group website) as HDF5 provides cross platform access, fast I/O, big data storage, and structured metadata. Once the archives are created, they are published to the Impact-Aware Robotics Archives Collection on the 4TU.ResearchData database. This database is therefore the storage space of the archives.

The front-end of the database is accessible through https://impact-aware-robotics-database.tue.nl/, which provides a search interface for users to select specific archives. Besides, the front-end provides direct information on the Objects, Environments, and Robots used within the recorded experiments of the archives.

More background information and details about the archives can be found in Deliverable D1.1 (will download file). More information about the project and motivation can be found on the project webpage of the I-AM project.

Table of content

- Overview
- Installation
- Usage of the scripts
- Contact

Overview

Overall description of the setup and the collected data

Installation

The code of this repository is all written in MATLAB and can directly be pulled from this repository.

If you are working on Linux, you need to download ffmpeg (for converting .bag files to .mp4 files). Follow these steps:\

```
sudo apt install ffmpeg
```

Figure 4: Screenshot of the public Gitlab README file with the instruction on how to use the data storage of the I.AM. dataset.

3.4. Adding Datasets to the Impact-Aware Robotics Database

The current version of the database contains datasets that are collected during the I.AM. project. How- ever, the database is intended to continue and grow beyond this project and to provide a platform for data storage for any experiment related to impact-aware robotics. To this end, we stimulate researchers in the field of impact-aware robotics to contribute to the database by providing the option to upload datasets to the database. In Sections 3.2 and 3.3, we discussed the open software we have created that can be used to collect and store the experimental data in a predefined structure used for the Impact- Aware Robotics Database. This software provides the end-user with the tools to create a dataset that can be uploaded to the database. At this stage, the procedure to do so is as follows, but the reader is referred to the database website [2] for the latest information about contributing to the database.

To upload a dataset, the end-user must create an account on 4TU.ResearchData [4]. This can be done either through an institutional login (if the institution the end-user is connected to is a partner of 4TU) or via an eduID[2], which can be created using an email address and name. End-users that are not connected to an institution that is partnered with 4TU can get up to 5GB of free storage, and after that, they must pay 4.5 euro per gigabyte[3]. End-users that are connected to a partner institution can upload more data, depending on the partner status (either gold, silver, or bronze member). Datasets can then be uploaded under different licenses and linked to the Impact-Aware Robotics Archive Collection [3] on 4TU.ResearchData by adding a metadata tag pointing to this collection. All datasets currently stored in the database are publicly available under the Creative Commons Attribution-NonCommercial-ShareAlike
4.0 International (CC-BY-NC-SA 4.0) license. We recommend using this license, but authors are free to choose differently. Best practice is to contact a TU/e contact mentioned on the website [2] for advice or support on data conversion and upload. After the dataset is uploaded, a data curator from the TU/e evaluates the dataset before it is added to the collection. Once the dataset is part of the collection on 4TU, it is automatically displayed on the website of the Impact-Aware Robotics Database.

[2] An eduID is an account that allows to login into several services connected to Dutch educational and research institutions, meant for users that are not connected to any of those institutions. More info can be found at `https://eduid.nl/en/`.

[3] More information can be found at 4TU.ResearchData, see `https://data.4tu.nl/info/en/about-4turesearchdata/membership-prospectus`.

4. DATA COLLECTION PLAN AND COLLECTED DATASETS

In Deliverable D1.1 and D1.4, the following expectations regarding the data collection plan were stated.

Responsible Partner:

- TOSS and BOX scenarios, mainly TU/e, with input from TUM, FRANKA EMIKA, CNRS, and EPFL.
- GRAB scenario: mainly EPFL and CNRS for providing equipment (dual arm KUKA robot, dual Panda arms, sensor I/O logging), support for hardware and acquisition software, performing of experiments; TU/e for setting up experimental conditions and storage of collected data in agreed format with inputs from TUM, CNRS, and EFPL.

Expected recording time:

- First year: TOSS/BOX
- Second year: BOX/GRAB
- Third year: GRAB

The current version of the Impact-Aware Robotics Database contains a total of 23 datasets related to all three different scenarios as a continuous effort of all partners in the project.
That is an increase with 7 additional datasets already, since the last publication in D1.4 December 2022. The database is still growing and a complete overview of the (current) uploaded datasets is shown in Table 1.
Note that this table shows the entire uploaded datasets as of 30 April 2023, so including those that were already published in D1.4 (December 2022), updated and added datasets are depicted in bold.

Nr.	Name	Creation Date	Size	Scenario
1	First version of an Impact Aware Manipulation (I.AM.) archive	25/09/2020	1.9GB	TOSS
2	Impact Aware Manipulation (I.AM.) archive containing suction cup recordings	18/11/2021	1.96GB	TOSS
3	Impact Aware Manipulation (I.AM.) archive containing box impact recordings	03/12/2021	2.48 GB	TOSS
4	Impact Aware Manipulation (I.AM.) archive containing suction cup release experiments	22/08/2022	1.02GB	TOSS
5	Impact Aware Manipulation (I.AM.) archive containing box-drop experiments for Parameter Identification of Box006	07/09/2022	1.95GB	TOSS

6	I.AM. archive containing box-drop experiments for Velocity Based Parameter Identification of Box007	24/10/2022	2.46GB	TOSS
7	I.AM. archive containing box-drop experiments for Velocity Based Parameter Identification of Box009	24/10/2022	2.41GB	TOSS
8	I.AM. archive containing box-drop experiments for Trajectory Based Parameter Identification of Box006	24/10/2022	1.1GB	TOSS
9	I.AM. archive containing box-toss experiments with Box005 for repeatability analysis	24/10/2022	3.73GB	TOSS
10	I.AM. archive containing box-toss experiments with Box005 for sensitivity analysis	25/10/2022	6.78GB	TOSS
11	I.AM. archive containing long range box-toss experiments with Box005 and Box006 for validation of impact models	26/10/2022	2.75GB	TOSS
12	Archive containing experiments with a drone impacting a rigid wall with the purpose of parameter identification	26/10/2022	821.02MB	-
13	I.AM. archive containing long range box-toss experiments with Box006, Box007, and Box009 for validation of impact aware object tracking	26/10/2022 **->** ***updated 30/01/2023***	308.44MB **->** ***updated: 357MB***	TOSS
14	I.AM. archive containing dual-arm grabbing experiments	14/11/2022	115MB	GRAB
15	I.AM. archive containing experiments of stacking boxes in totes, relevant for the BOX scenario.	01/12/2022	947MB	BOX
16	I.AM. archive containing impact experiments with a Franka Emika Panda robot	01/12/2022	452MB	GRAB
17	I.AM. archive containing long-range tossing experiments for Trajectory Based Parameter Identification of Box012	04/01/2023	379MB	TOSS
18	**I.AM. archive containing long-range tossing experiments for Trajectory Based Parameter Identification of Box009**	**04/01/2023**	**392MB**	**TOSS**

19	I.AM. archive containing long-range tossing experiments for Trajectory Based Parameter Identification of Box007	04/01/2023	410MB	TOSS
20	I.AM. archive containing long-range tossing experiments for Trajectory Based Parameter Identification of Box004	04/01/2023	400MB	TOSS
21	I.AM. archive containing long range box-toss experiments with Box012 for validation of impact models	04/01/2023	395MB	TOSS
22	I.AM. archive containing long range box-toss experiments with Box009 for validation of impact models	04/01/2023	391MB	TOSS
23	I.AM. archive containing long range box-toss experiments with Box007 for validation of impact models	04/01/2023	410MB	TOSS
24	I.AM. archive containing long range box-toss experiments with Box004 for validation of impact models	04/01/2023	406MB	TOSS

Table 1: Overview of the uploaded datasets to the Impact-Aware Robotics database, updated on 30/04/2023.

5. CONCLUSION

This deliverable D6.6 provides the updated description of the I.AM. repository structure and the implementation towards the final I.AM. database that is being released till end of 2023.
It provides mainly an update over D1.4 (December 2022) with increased datasets as well as new public Gitlab repos available with descriptions on how to contribute to the database.

6. REFERENCES

[1] M. J. Jongeneel, S. Dingemans, and A.Saccon, "The Impact-Aware Robotics Database," Pre-print on HAL, Nov. 2022. [Online]. Available: https://hal.science/hal-03900923

[2] A. Saccon and M. J. Jongeneel, "TU/e Impact-Aware Robotics Database," Eindhoven, The Nether- lands, 2022, [Accessed 10-Oct-2022]. [Online]. Available: https://impact-aware-robotics-database. tue.nl/

[3] ——, "Impact-Aware Robotics Archives Collection," 4TU.ResearchData, Apr. 2021, Dataset. [Online]. Available: https://doi.org/10.4121/c.5405187

[4] The 4TU.Federation, "4TU.ResearchData data and software repository," 2008-2022. [Online]. Available: https://data.4tu.nl/info/en/

[5] M. J. Jongeneel, J. d. Ouden, and A. Saccon, "D1.4, Publication of I.AM. Dataset (update 1.1)," 2020. [Online]. Available: https://www.i-am-project.eu/index.php/results/deliverables