





<h2>Datacard 3</h2>		<p>V2AIX: A Multi-Modal Real-World Dataset of ETSI ITS V2X Messages in Public Road Traffic</p>	
<p>DATASET LINK</p>		<p>DATA PUBLISHER(S)</p>	
<p><i>Provide a link to the dataset:</i></p>		<p><i>Companies or people who created the dataset</i></p>	
<p>https://v2aix.ika.rwth-aachen.de/</p>		<p>Guido Küppers, guido.kueppers@ika.rwth-aachen.de</p> <p>Jean-Pierre Busch, jean-pierre.busch@ika.rwth-aachen.de</p> <p>Lennart Reiher, lennart.reiher@ika.rwth-aachen.de</p> <p>Lutz Eckstein</p>	
<h3>Funding Sources (Optional)</h3>			
<p>INSTITUTION(S)</p>		<p>FUNDING OR GRANT SUMMARY(IES)</p>	
<p><i>Provide the names of the funding institution(s):</i></p>		<p><i>Add this section if the dataset was generated with AITHENA funding</i></p>	
 <p>Funded by the European Union</p>		<p>The Althena project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number 101076754. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.</p>	
 <p>Federal Ministry of Education and Research</p>		<p>This work is additionally accomplished within the projects 6GEM and AUTotech.agil. We acknowledge the financial support for the projects by the Federal Ministry of Education and Research of Germany (BMBF) for 6GEM (FKZ 16KISK036K) and AUTotech.agil (FKZ 01IS22088A).</p>	

Dataset Overview

DATA SUBJECT(S)

CONTENT DESCRIPTION

Select **ALL APPLICABLE** subjects contained the dataset:
[Definition of sensitive data according to GDPR](#)

Provide a short description of the content of the data and the motivation. This section is open for any description you consider helpful. Ex: what problems you were looking to solve

- Sensitive Data about people
- Non-Sensitive Data about people
- Data about natural phenomena
- Data about places and objects
- Synthetically generated data
- Data about systems or products and their behaviours
- Unknown
- Others (Please specify):

V2AIX is a multi-modal real-world dataset of ETSI ITS messages gathered in public road traffic. The data were collected in measurement drives and with stationary infrastructure, including more than 230 000 V2X messages from more than 1800 vehicles and roadside units in public road traffic.

The full dataset is publicly available for non-commercial use at v2aix.ika.rwth-aachen.de

DATASET SNAPSHOT

Provide a snapshot of the dataset:

(Use the additional notes to include relevant information, considerations, and links to table(s) with more detailed breakdowns.) The format of how the data is presented is not important, the following are examples

Size of Dataset	1257 GB (ROSbags), 20 GB (JSON)
Number of Instances (Ex: images or sequences)	
Labelled Classes (List of classes)	No labelled data
Number of Labels (Ex: number of total objects present in dataset)	No labelled data
Average Labels Per Instance (Ex: avg number of objects per image)	No labelled data
Algorithmic Labels (Automatically labeled)	No labelled data
Human Labels (Labelled by human annotators)	No labelled data
Other Characteristics	1851 ITS Stations 215656 Cooperative Awareness Messages (CAMs) 994 Decentralized Environmental Notification Messages (DENMs) 2314 Map Extended Messages (MAPEM)

18192 Signal Phase and Timing Extended Messages (SPATEM)

23h of V2X Duration

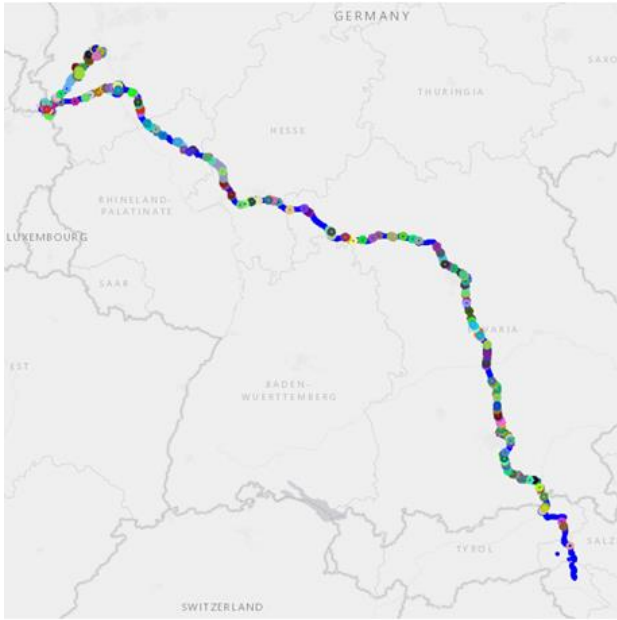
1420 km V2X track distance

Scene Descriptors and ODD

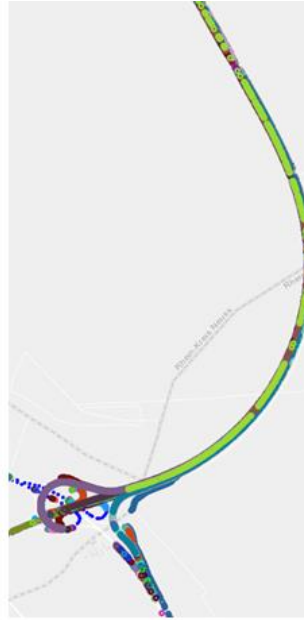
Provide descriptors of the scene that was recorded taking the following parameters and add more as needed

Labels and percentage (if applicable)

- *Weather*
 - *Sunny*
 - *Cloudy*
 - *Partly cloudy/sunny*
- *Daylight*
 - *Day*
 - *Sunset/Sunrise*
- *Environment*
 - *Urban*
 - *Suburban*
 - *Rural*
 - *Highway*



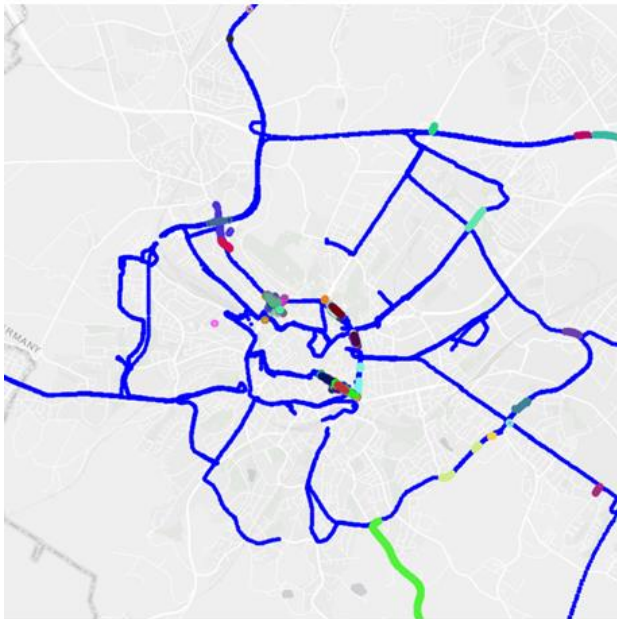
Overview



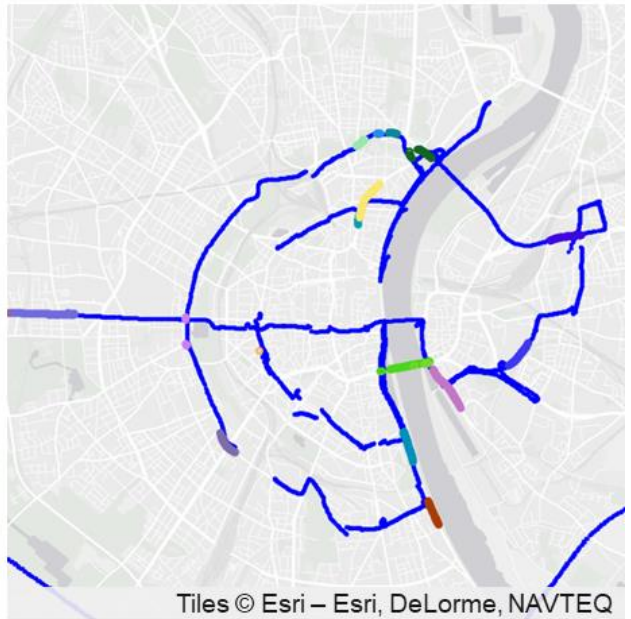
Highway A44



Dusseldorf



Aachen



Cologne

Tiles © Esri – Esri, DeLorme, NAVTEQ

Blue line indicating the driven trajectory during recording, coloured dots showing ITS stations transmitting CAMs.

Sensitivity of Data

SENSITIVITY TYPE(S)	FIELD(S) WITH SENSITIVE DATA	SECURITY AND PRIVACY HANDLING
<p>Select all applicable data types present in the dataset:</p>	<p><i>List fields in the dataset that contain S/PII, and specify if their collection was intentional or unintentional.</i></p> <p><i>Use additional notes to capture any other relevant information or considerations.</i></p>	<p><i>Summarize the measures or steps to handle sensitive data in this dataset.</i></p> <p><i>Use additional notes to capture any other relevant information or considerations.</i></p>
<p>S/PII</p> <p>Anonymous Data</p> <p>Traffic Data</p>	<p>Unintentionally Collected Sensitive Data</p> <p>The dataset contains camera-images which could be used to infer license-plate information from specific vehicles.</p>	<p>The camera images were processed prior to publication of the data set to blur license plates (and theoretically also faces) to make them unrecognizable.</p>

MAINTENANCE STATUS	VERSION DETAILS	MAINTENANCE PLAN
<p>Select one:</p>	<p>Provide details about this version of the dataset:</p>	<p>Summarize the maintenance plan for the dataset:</p> <p>If using a tool for the versioning, you can mention here which one and how it works/what it records...</p>
<p>Regularly Updated / Actively Maintained</p>	<p>Current Version: 1.0</p> <p>Last Updated: 03/2024</p> <p>Release Date: 03/2024</p>	<p>Versioning: None</p> <p>Updates: None</p> <p>Errors: None</p> <p>Feedback: None</p> <p>Additional Notes: None</p>

Example of Data Points

PRIMARY DATA MODALITY	EXAMPLE OF DATAPOINT
<i>Select one:</i>	<i>List the fields in data points and their descriptions. If you are using a standardised data format you can mention it and provide a link instead of a description</i> <i>(Usage Note: Describe each field in a datapoint. Optionally use this to show the example.)</i>
Image Data Pointcloud Data Time Series Geospatial Data	Description of data can be found here: https://v2aix.ika.rwth-aachen.de/Usage_Instructions

Access, Retention, & Wipeout

Access

ACCESS TYPE

DOCUMENTATION LINK(S)

Select **one**:

Provide links that describe documentation to access this dataset:

External - Open Access

<https://v2aix.ika.rwth-aachen.de/>

https://github.com/ika-rwth-aachen/etsi_its_messages

Provenance

Collection

METHODOLOGY DETAIL(S)	SOURCE DESCRIPTION(S)	DATA PROCESSING
<p><i>Provide a description of each collection method used.</i></p> <p><i>Use additional notes to capture any other relevant information or considerations.</i></p> <p><i>(Usage Note: Duplicate and complete the following for collection method type.)</i></p>	<p><i>Provide a description of the source of data.</i></p> <p><i>It can be software for synthetic data or the description of the vehicle...</i></p> <p><i>(Usage Note: Duplicate and complete the following for collection method type.)</i></p>	<p><i>Summarize how data from different sources or methods aggregated, processed, or connected.</i></p> <p><i>(Usage Note: Duplicate and complete the following for each source OR collection method.)</i></p>
<p>Data collection methodology is described in the given paper: V2AIX: A Multi-Modal Real-World Dataset of ETSI ITS V2X Messages in Public Road Traffic (also available on arXiv)</p>	<p>Data collection source/vehicle is described in the given paper: V2AIX: A Multi-Modal Real-World Dataset of ETSI ITS V2X Messages in Public Road Traffic (also available on arXiv)</p>	<p>Data processing steps are described in the given paper: V2AIX: A Multi-Modal Real-World Dataset of ETSI ITS V2X Messages in Public Road Traffic (also available on arXiv)</p>
<p>COLLECTION PROCESS</p>		
<p><i>Provide a description or a link to a document that describes the criteria for the recordings or generation of data</i></p>		
<p>Data collection process is described in the given paper: V2AIX: A Multi-Modal Real-World Dataset of ETSI ITS V2X Messages in Public Road Traffic (also available on arXiv)</p>		

Human and Other Sensitive Attributes

SENSITIVE HUMAN ATTRIBUTE(S)

Select **all attributes** that are represented (directly or indirectly) in the dataset.

[Note: there are demonstrative examples of potentially relevant signals for a given dataset analysis]

Gender

Socio-economic status

Geography

Language

Age

Culture

Others (Please Specify):

REPRESENTED GROUPS OF PEOPLE

No matter if you record it as part of your data/labels. If you are recording people, we would like to know what groups of people are present in your data. If possible, provide a distribution

Non-disabled adult

Non-disabled child

Non-disabled elderly

Disabled adult

Disabled child

Disabled elderly



The [Data Cards Playbook](#) by Google Research is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

You are free to share and adapt this work under the [appropriate license terms](#).