

团 体 标 准

T/SHV2X 10—2025

汽车驾驶自动化系统通用语料库 第 2 部分：术语和定义

General corpus for automotive driving automation system—
Part 2: Terms and Definitions

2025 - 07 - 25 发布

2025 - 07 - 25 实施

全国团体标准信息平台

目 次

前 言.....	II
引 言.....	III
1 范围.....	1
2 规范性引用文件.....	1
3 术语.....	1
参 考 文 献.....	4

中国团体标准

前 言

《汽车驾驶自动化系统通用语料库》系列标准拟分为六个部分：

- 第1部分：总体要求；
- 第2部分：术语和定义；
- 第3部分：语料数据采集；
- 第4部分：语料数据清洗；
- 第5部分：语料数据标注；
- 第6部分：语料数据测试。

本文件为第2部分。

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由上海市车联网协会提出并归口。

本文件起草单位：上海库帕思科技有限公司、智能汽车创新发展平台（上海）有限公司、智己汽车科技有限公司、地平线征程（上海）科技有限公司、上海临港绝影智能科技有限公司、上海机动车检测认证技术研究中心有限公司、上海金桥智能网联汽车发展有限公司、北京一辅智行科技有限公司、上海优咔网络科技有限公司、亿咖通（上海）技术有限公司、滴水智行科技有限公司、上海航盛实业有限公司、上海阶跃星辰智能科技有限公司、东华大学、上海交通大学、上研智联智能出行科技（上海）有限公司、北京赛目科技股份有限公司、苏州柏川数据科技有限公司、华为技术有限公司。

本文件主要起草人：山栋明、黄海清、施佳樑、郭辉、赵九花、贺锦鹏、蒋达夫、周剑鸣、黄剑其、邵亚萌、徐春雷、解瀚光、朱雷、谭龙欢、于峰、滕添益、张帆、林瑜、周轶、袁月明、李学根、张裕珍、曹宇、邓思文、贺仁驹、李晨歌、马昊、陈紫娟、杨闻博、丁楚晨、李想、蔡雨辰、黄鹏飞、刘建业、李勋宏、马骏、李轶刚、刘壹青、田浩、郭晓宾、董连飞、范昌琪、李璟、孙雯、陈巧慧、王娜、沈滨、孔令和、鲁江东、杨伟利、林翔、何丰、谭哲、薛晓卿、刘兴、马东升、刘鹏宇、邓子涵。

本文件首批承诺执行单位：上海库帕思科技有限公司、智能汽车创新发展平台（上海）有限公司、智己汽车科技有限公司、地平线征程（上海）科技有限公司、上海临港绝影智能科技有限公司、上海机动车检测认证技术研究中心有限公司、上海金桥智能网联汽车发展有限公司、北京一辅智行科技有限公司、上海优咔网络科技有限公司、亿咖通（上海）技术有限公司、滴水智行科技有限公司、上海航盛实业有限公司、上海阶跃星辰智能科技有限公司、东华大学、上海交通大学、上研智联智能出行科技（上海）有限公司、北京赛目科技股份有限公司、苏州柏川数据科技有限公司、华为技术有限公司。

引 言

为了实现可靠的汽车驾驶自动化,大量准确的语料数据至关重要。语料数据是自动驾驶系统的基础,它能够帮助系统理解复杂的交通环境、做出明智的决策,并不断优化性能。随着端到端模型整合度越来越高,对训练语料的规模和质量要求也成倍增加。大规模、高质量的数据标注(特别是端到端感知数据和推理数据)是构建安全、可靠、高性能自动驾驶系统的基石。围绕“采、洗、标、测”一体化流程,制定《汽车驾驶自动化系统通用语料库》系列标准,包括:

- 第1部分:总体要求,明确系列标准建设的总体要求,为后续各部分标准的制定提供指导和基础;
- 第2部分:术语和定义,统一系列标准建设过程中涉及的关键术语和定义;
- 第3部分:语料数据采集,对语料数据资源格式提出规范要求;
- 第4部分:语料数据清洗,针对采集好的数据,对语料数据清洗的流程与质量要求提出规范要求;
- 第5部分:语料数据标注,针对清洗好的数据,对语料数据标注的适用场景、标注内容、标注方式、数据存储提出规范要求;
- 第6部分:语料数据测试,针对标注好的数据,对语料数据测试的流程和质量要求提出规范要求。

本文件为第2部分术语和定义。通过本文件的制定,对《汽车驾驶自动化系统通用语料库》系列标准中所涉及的术语部分进行统一规范。通过构建可复用的标准化自动驾驶训练数据集,促进汽车驾驶自动化系统语料资源高效流通利用,降低业内企业重复标注、处理和管理数据等方面的巨额成本,提升汽车企业竞争力且推进汽车产业健康发展。

汽车驾驶自动化系统通用语料库 第2部分：术语和定义

1 范围

本文件规定了汽车驾驶自动化系统通用语料库建设过程中涉及的主要术语和定义，为语料库的设计、采集、清洗、标注、测试提供统一的语言基础。

本文件适用于企业、研究机构对汽车驾驶自动化系统通用语料库的研究、开发、维护、应用、评估等工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB 5768.2—2009 道路交通标志和标线 第2部分：道路交通标志

GB 5768.3—2022 道路交通标志和标线 第3部分：道路交通标线

GB/T 25000.12—2017 系统与软件工程 系统与软件质量要求和评价(SQuaRE)第12部分：数据质量模型

GB/T 36073—2018 数据管理能力成熟度评估模型

GB/T 40429—2021 汽车驾驶自动化分级

GB/T 41871—2022 信息安全技术 汽车数据处理安全要求

GB/T 42755—2023 人工智能面向机器学习的数据标注规程

GB/T 44373—2024 智能网联汽车 术语和定义

CH/T 8023—2011 机载激光雷达数据处理技术规范

3 术语

3.1

驾驶自动化系统 driving automation system

由实现驾驶自动化的硬件和软件所共同组成的系统。

[来源：GB/T 40429—2021, 2.2]

3.2

数据标准 data standard

数据的命名、定义、结构和取值的规则。

[来源：GB/T 36073—2018, 3.7]

3.3

语料数据 corpus data

构成语料库的原始或处理后的数据，包括但不限于传感器数据、驾驶行为数据、环境信息等。

3.4

语料数据资源 corpus data resources

以电子化形式记录和保存的具备原始性、可机器读取、可供社会化再利用的数据集合。

3.5

数据表征模式 data representation modes

计算机中存储和表达数据的形式或结构。

3.6

传感器数据 sensor data

由车辆搭载的各类传感器（如摄像头、雷达、激光雷达、惯性测量单元等）采集的原始信息。

3.7

摄像头 camera

通过感光器件获取图像信息的装置。

[来源：GB/T 44373—2024, 5.4]

3.8

激光雷达 lidar

利用激光束探测目标并获取目标信息的装置。

[来源：GB/T 44373—2024, 5.5]

3.9

毫米波雷达 millimeter wave radar

利用毫米波波段的电磁波发现目标并获取目标信息的装置。

[来源：GB/T 44373—2024, 5.6]

3.10

超声波传感器 ultrasonic sensor

利用超声波探测目标并获取目标信息的装置。

[来源：GB/T 44373—2024, 5.7]

3.11

数据采集 data collection

按照预定方案，通过各种传感器和设备获取汽车驾驶自动化系统相关数据的过程。

3.12

数据清洗 data cleaning

对采集到的原始数据进行筛选、去噪、校正和格式化等处理，提升数据质量的过程。

3.13

数据标注 data labeling

给数据样本指定目标变量和赋值的过程。

[来源：GB/T 42755—2023, 3.1]

3.14

数据批次 data batch

同一采集任务中，具有连续时空关联性的一组多模态数据集合。

3.15

数据抽帧 data frame sampling

从连续的传感器数据流中按特定规则选取子集。

3.16

时序对齐 temporal alignment

将不同传感器在同一时刻采集的数据进行时间戳校准和同步的过程。

- 3.17
数据质量 data quality
在指定条件下使用时,数据的特性满足明确的和隐含的要求的程度。
[来源:GB/T 25000.12—2017,4.3]
- 3.18
数据安全 data security
数据的机密性、完成性和可用性。
[来源:GB/T 36073—2018,3.11]
- 3.19
极端情况 corner case
在自动驾驶场景中出现频率低但对系统安全性影响重大的边缘情况,如恶劣天气、罕见交通状况、传感器故障等。
- 3.20
交通参与者 traffic participants
交通流中的行人、交通工具和动物等。
- 3.21
道路交通信号灯 road traffic signal lamps
交通流中指挥交通运行的信号灯。
- 3.22
道路交通标志 road traffic signs
以颜色、形状、字符、图形等向道路使用者传递信息,用于管理交通的设施。
[来源:GB 5768.2—2022]
- 3.23
道路交通标线 road traffic markings
道路上的各种线条、箭头、文字、图案及立面标记、实体标记、突起路标和轮廓标等所构成的交通设施。
[来源:GB 5768.3—2009]
- 3.24
可行驶区域 drivable area
交通参与者可以通行的道路区域。
- 3.25
场景采集环境 scenario acquisition environment
场景中的光源信息、路面材质和天气情况等。
- 3.26
高精地图 high definition map
相比导航地图,能提供精度更高、内容更丰富的道路拓扑、拓扑关系、位置、几何、交通标识、交通信号设施等地图属性,为智能网联汽车提供环境信息的地图。
[来源:GB/T 44373—2024,5.8]

参 考 文 献

- [1] T/CSAE 212—2021 智能网联汽车场景数据图像标注要求与方法
 - [2] T/SAIAS 015—2024 语料库建设导则
 - [3] T/SHV2X 9—2025 汽车驾驶自动化系统通用语料库 第1部分：总体要求
 - [4] T/SHV2X 1—2025 汽车驾驶自动化系统通用语料库 第3部分：语料数据采集
 - [5] T/SHV2X 11—2025 汽车驾驶自动化系统通用语料库 第4部分：语料数据清洗
 - [6] T/SHV2X 12—2025 汽车驾驶自动化系统通用语料库 第5部分：语料数据标注
 - [7] 智能网联汽车时空数据安全处理基本要求
 - [8] 自然资发〔2024〕139号 自然资源部关于加强智能网联汽车有关测绘地理信息安全管理的通知
-