
archi-intelligence Research Series

Working Paper 2026-01

架构的世纪迁徙

从工程本体论到具身智能的统一图景

跨产业架构演进、跨域基准与成熟度框架的学术综述

archi-intelligence Research Team

2026 年 6 月发布

archi-intelligence.org

0.0.1 F.3 版权与许可页

版权信息 / Copyright Notice

本文档 © 2026 archi-intelligence Research Team

本作品在 Creative Commons Attribution 4.0 International (CC-BY 4.0) 许可下发布。您可以自由：

- 分享—在任何媒介以任何形式复制和重新分发本作品
- 改编—修改、转换或基于本作品进行二次创作

唯一条件：必须以适当方式标明出处、提供许可链接、说明是否做出了修改。

License URL: <https://creativecommons.org/licenses/by/4.0/>

数字对象唯一标识符 / DOI

DOI: 10.5281/zenodo.20357858 Permanent Archive: <https://zenodo.org/records/20357858>

引用格式建议 / Suggested Citation

archi-intelligence Research Team. (2026). The Architectural Migration of the Century: From Engineering Ontology to Embodied Intelligence (Working Paper 2026-01). archi-intelligence Research Series. <https://doi.org/10.5281/zenodo.20357858>

BibTeX 格式

```
@techreport{archi_intelligence_2026_01,  
  title      = {The Architectural Migration of the Century:  
                From Engineering Ontology to Embodied  
                Intelligence},  
  author     = {{archi-intelligence Research Team}},  
  institution = {archi-intelligence},  
  year       = {2026},  
  number     = {2026-01},  
  type       = {Working Paper},  
  series     = {archi-intelligence Research Series},  
  doi        = {10.5281/zenodo.20357858},  
  url        = {https://archi-intelligence.org/research/2026-01}  
}
```

联系方式 / Contact

- 研究团队 / Research Team: research@archi-intelligence.org
- 媒体咨询 / Press Inquiries: press@archi-intelligence.org
- 内容更正 / Corrections: corrections@archi-intelligence.org
- 网站 / Website: <https://archi-intelligence.org>

0.0.2 F.4 利益声明 (COI Disclosure)

利益冲突与编辑独立性声明

为保持本研究系列的学术独立性与公信力,archi-intelligence Research Team 在此明确以下事项:

1. 关于本研究机构

archi-intelligence 是致力于 Architecture Intelligence (架构智能, AI²) 研究范式建立的独立学术研究机构。本机构的使命是通过公开方法论、开放数据引用、严格同行评审, 推动跨产业架构工程实践的标准化与可比性。

2. 关于资金与赞助

本研究系列在 v1.0 发布阶段未接受任何被评估对象 (包括但不限于本报告中提及的 OEM、Tier-1 供应商、平台公司、芯片厂商) 的直接或间接资金支持。

archi-intelligence 在初始阶段由 Arkimind (一家专注于汽车电子电气架构智能化的商业实体) 提供 seed sponsorship。这一关系在本声明中明确披露, 并通过以下治理隔离机制保障编辑独立性:

- 编辑/研究团队不向 Arkimind 销售或业务团队汇报
- 研究人员的薪酬不与 Arkimind 任何商业指标挂钩
- 报告发布前, Arkimind 商业团队无审阅或修改权限
- 评估方法论与原始数据来源 100% 公开
- 任何被评估对象可提交修正反馈, 处理流程公开

3. 关于编辑独立性

本研究的方法论选择、案例评估、结论表述均由 archi-intelligence Research Team 独立做出。研究判断不受任何商业利益、政治立场、地缘政治偏好的影响。

我们承认这一编辑独立性最终需要由读者通过持续审视我们的研究质量来验证。我们欢迎来自学术界、产业界、监管机构的批评性反馈, 并承诺在后续版本中公开记录所有重要修正。

4. 关于数据来源与时间窗

本研究采用 2024 年 1 月至 2026 年 1 月期间的公开来源信息, 包括但不限于:

- 上市公司财报与 SEC/HKEX 公开文件
- OEM 与供应商官方技术发布 (AI Day、HDC、IAA、CES 等)
- 已公开的专利文献
- 同行评审学术论文与会议出版物
- 行业标准组织发布 (AUTOSAR、ISO、SAE、IEEE 等)
- 主流技术媒体报道与产业分析

对于无法独立验证的传闻信息或匿名来源信息, 本研究采取保守立场——要么不引用, 要么明确标注为”基于公开报道的方向性推断”。

5. 关于研究限制

本研究作为综述性工作, 明确承认数据可获得性的不对称性。Tesla、NVIDIA、Google 等公司的公开技术细节相对完整, 而华为、小米、部分中国 OEM 的车端底层实现细节公开度有限。本报告在比较时已力求标注这一不对称性, 但读者应警惕将”公开度差异”误读为”技术能力差异”。

6. 关于商标与商品名

本报告中提及的所有商标、商品名、产品代号均归各自所有者所有。引用这些名称的目的纯为学术讨论与产业分析, 不构成任何商业关联、背书或推广。

7. 关于免责

本报告不构成投资建议、商业咨询或任何形式的法律意见。基于本报告内容做出的任何商业决策, 由决策方独立承担风险与责任。

- 研究团队联系: research@archi-intelligence.org
- 更正与反馈: corrections@archi-intelligence.org
- 完整治理章程: <https://archi-intelligence.org/governance>

0.0.3 F.5 详细目录

Contents

0.0.1	F.3 版权与许可页	1
0.0.2	F.4 利益声明 (COI Disclosure)	2
0.0.3	F.5 详细目录	4
0.0.4	F.6 图表索引 (17 个图表完整清单)	6
0.0.5	F.7 缩略语表	7
1	架构的世纪迁徙：从工程本体论到具身智能的统一图景	10
1.1	摘要	10
1.2	引言：为什么”架构”是这个十年最被低估的元问题	10
1.3	第一章架构的本体论：从维特鲁威到 Architecture Intelligence	12
1.3.1	1.1 起源：希腊语 <i>archi-tekton</i> 与维特鲁威三原则	12
1.3.2	1.2 从模式语言到软件架构：20 世纪后半叶的概念跃迁	12
1.3.3	1.3 当代严谨定义：ISO 42010、TOGAF 与 Shaw-Garlan-Booch-Kruchten 谱系	13
1.3.4	1.4 架构学的两条底层定律	13
1.3.5	1.5 架构观的当代谱系：从结构到平台到生态	14
1.3.6	1.6 架构演进的元规律	14
1.3.7	1.7 引入 Architecture Intelligence	14
1.4	第二章汽车电子电气架构：从分布式 ECU 到中央计算的演进	16
1.4.1	2.1 概念边界：E/E 架构不只是线束与 ECU 拓扑	16
1.4.2	2.2 演进的物理压力：ECU 数量危机	16
1.4.3	2.3 博世六阶段演进框架	16
1.4.4	2.4 2025—2026 年主要玩家算力图谱	18
1.4.5	2.5 大众 CARIAD 困局与 Rivian 联盟：垂直闭环受挫到平台联盟的转型范例	19
1.4.6	2.6 参考框架与标准矩阵	20
1.5	第三章跨产业架构对标：五维矩阵、失效哲学与两层结构	22
1.5.1	3.1 五维基准矩阵	22
1.5.2	3.2 硬件趋同：全球架构基底的收敛	23
1.5.3	3.3 失效哲学：fail-soft vs fail-operational 的根本分水岭	23
1.5.4	3.4 Hypervisor 作为数字-物理鸿沟的桥梁	24
1.5.5	3.5 数据/AI 架构的边-云分工：双向收敛	24
1.5.6	3.6 生态：三种集结哲学	25
1.5.7	3.7 两层结构：基础设施趋同，控制语义与证明义务分化	25

1.6	第四章三类架构原型的深度基准	27
1.6.1	4.1 垂直闭环型: Tesla 的形态无关感知栈	27
1.6.2	4.2 跨设备生态型: 华为与小米的两条中国路径	29
1.6.3	4.3 平台赋能型: Google 的联邦 AI 平台与 Waymo 的可证明安全架构	31
1.6.4	4.4 平台赋能型的其他亚类	34
1.6.5	4.5 三类原型的横向对比矩阵	37
1.7	第五章架构成熟度的统一框架: AR0—AR5 与 AI ² -ML	43
1.7.1	5.1 启发与方法论	43
1.7.2	5.2 AR0—AR5 架构能力门槛框架	43
1.7.3	5.3 AR 分级的可视化	44
1.7.4	5.4 引入 Architecture Intelligence Maturity Levels (AI ² -ML)	45
1.7.5	5.5 AR 与 SAE J3016 的关系	48
1.8	第六章总结、研究边界与开放问题	50
1.8.1	6.1 核心论点综述	50
1.8.2	6.2 七个架构控制点	50
1.8.3	6.3 收敛-分化的终局图景	51
1.8.4	6.4 研究边界与限制	53
1.8.5	6.5 开放问题	53
1.8.6	6.6 致谢与方法论声明	54
1.8.7	6.7 致研究界与产业界的开放邀请	54
2	Back Matter 附录	59
2.0.1	附录 A: 术语表 Glossary	60
2.0.2	附录 B: AR & AI ² -ML 评分快速参考卡	63
2.0.3	附录 C: 跨产业 5 维对标完整数据表	65
2.0.4	附录 D: archi-intelligence Research Series 简介	70
3	编辑修订指南	73

0.0.4 F.6 图表索引 (17 个图表完整清单)

图表 / Figures

- Figure 1.1 两千年架构概念演化时间轴 (Two-Thousand-Year Architectural Concept Timeline)
- Figure 1.2 ISO/IEC/IEEE 42010 四元概念模型
- Figure 1.3 架构演化的元规律
- Figure 2.1 博世六阶段 E/E 架构演进图 (Bosch' s Six-Stage E/E Architecture Evolution)
- Figure 2.2 2025-2026 汽车 SoC 算力图谱
- Figure 2.3 CARIAD E3 架构代际困境与转型路径
- Figure 3.1 五大产业 × 五维度交叉对标矩阵
- Figure 3.2 双层趋同-分化模型图
- Figure 4.1 三类架构原型拓扑对比
- Figure 4.2 Tesla 跨形态复用机理图
- Figure 4.3 华为 vs 小米架构哲学对比
- Figure 4.4 Google + Waymo 平台赋能拓扑
- Figure 4.5 三类原型横向对比矩阵 [新增]
- Figure 5.1 AR0-AR5 能力门槛阶梯图
- Figure 5.2 AR × AI²-ML 二维评估视图 (旗舰图)
- Figure 5.3 跨行业” 工具-系统” 匹配模式对比
- Figure 6.1 七大架构控制点网络图
- Figure 6.2 收敛-分化终局图景可视化

表格 / Tables

- Table 1.1 Kruchten 4+1 视图模型
- Table 2.1 博世 E/E 架构六阶段对照表
- Table 2.2 主要玩家 SoC 算力规格对比
- Table 3.1 跨产业五维基准矩阵
- Table 4.1 华为与小米架构对比
- Table 4.2 三类原型横向对比矩阵 [新增]
- Table 5.1 AR0-AR5 阶段定义对照
- Table 5.2 AI²-ML 五级定义与典型工具
- Table 6.1 七个架构控制点与典型持有者

0.0.5 F.7 缩略语表

缩略	全称	中文
AD	Autonomous Driving	自动驾驶
ADAS	Advanced Driver-Assistance Systems	高级驾驶辅助系统
ADR	Architecture Decision Record	架构决策记录
ADS	Autonomous Driving System (华为)	华为自动驾驶系统
AI ²	Architecture Intelligence	架构智能
AI ² -ML	Architecture Intelligence Maturity Levels	架构智能成熟度
AOSP	Android Open Source Project	Android 开源项目
API	Application Programming Interface	应用程序接口
AR	Architecture Readiness (本研究框架)	架构成熟度
ASIL	Automotive Safety Integrity Level	汽车安全完整性等级
AUTOSAR	AUTomotive Open System ARchitecture	汽车开放系统架构
BEV	Bird's-Eye View / Battery Electric Vehicle	鸟瞰图 / 纯电动汽车
CAN	Controller Area Network	控制器局域网
CARIAD	Car.I.AM Digital	大众数字子公司
CEA	China Electronic Architecture (大众中国)	大众中国电子架构
CNCF	Cloud Native Computing Foundation	云原生计算基金会
COI	Conflict of Interest	利益冲突
DCU	Domain Control Unit	域控制器
DDS	Data Distribution Service	数据分发服务
DoF	Degree of Freedom	自由度

缩略	全称	中文
E/E	Electrical/Electronic	电子电气
ECU	Electronic Control Unit	电子控制单元
EEA	Electrical/Electronic Architecture	电子电气架构
FSD	Full Self-Driving (Tesla)	特斯拉完全自动驾驶
FuSa	Functional Safety	功能安全
GDPR	General Data Protection Regulation	通用数据保护条例
HAL	Hardware Abstraction Layer	硬件抽象层
HDC	Huawei Developer Conference	华为开发者大会
HIMA	Huawei Intelligent Mobility Alliance	鸿蒙智行
HMI	Human-Machine Interface	人机界面
HPC	High-Performance Computing	高性能计算
LLM	Large Language Model	大语言模型
MCU	Microcontroller Unit	微控制器
MDC	Mobile Data Center (华为)	移动数据中心
NN	Neural Network	神经网络
NPU	Neural Processing Unit	神经处理单元
OEM	Original Equipment Manufacturer	整车制造商
OS	Operating System	操作系统
OTA	Over-the-Air (Update)	空中下载升级
PIPL	Personal Information Protection Law (China)	个人信息保护法
ROS	Robot Operating System	机器人操作系统
RTOS	Real-Time Operating System	实时操作系统
SDV	Software-Defined Vehicle	软件定义汽车
SoC	System on a Chip	系统级芯片

缩略	全称	中文
SOA	Service-Oriented Architecture	面向服务的架构
SOTIF	Safety Of The Intended Functionality	预期功能安全
SSP	Scalable Systems Platform (VW Group)	大众可扩展系统平台
TEE	Trusted Execution Environment	可信执行环境
TFLOPS	Trillion Floating-point Operations per Second	万亿次浮点运算/秒
TOPS	Tera Operations per Second	万亿次运算/秒
V2X	Vehicle-to-Everything Communication	车与万物通信
VLA	Vision-Language-Action (Model)	视觉-语言-动作模型
WEWA	World Engine + World Action Model (华为 ADS 4.0)	华为 ADS 4.0 架构
ZCU	Zone Control Unit	区域控制器

1 架构的世纪迁徙：从工程本体论到具身智能的统一图景

跨产业架构演进、跨域基准与成熟度框架的学术综述

Architecture Intelligence Research Series · Working Paper 2026-01

1.1 摘要

本文以“架构”这一跨越两千年的工程概念为起点，系统梳理了从维特鲁威三原则到 ISO/IEC/IEEE 42010 架构描述本体论的概念史，并沿汽车电子电气 (E/E) 架构、智能手机/消费电子、互联网/云、机器人、具身人工智能五大产业剖面，建立了一个统一的横向基准框架。研究在**硬件、软件、数据/AI、生态、安全**五个维度上完成跨域对标，识别出三类架构原型——**垂直闭环型** (Tesla)、**跨设备生态型** (华为、小米) 与 **平台赋能型** (Google/Waymo、NVIDIA、Mobileye、Bosch、Aptiv)——并对其内在哲学差异、复用机理与限制条件进行了详细剖析。

研究的核心判断是：**当代各产业的架构正在经历“基础设施趋同、控制语义与证明义务分化”的双层演化**。在共享底座 (异构 SoC、虚拟化、服务化接口、模型训练管线、OTA 流水线、数字孪生) 上，汽车与机器人将长期保留独立的失效哲学、安全壳与责任归属层；消费电子与云则继续以 fail-soft 与快速恢复为主导。本文进一步提出 **AR0—AR5 架构能力门槛框架**，类比 SAE J3016 自动驾驶分级，但明确强调“能力门槛而非时间线，低阶不会消失而会与高阶长期共存”。在此基础上，本文引入 **Architecture Intelligence (架构智能, AI²)** 作为统摄性研究范式，并提出 **AI²-ML (Architecture Intelligence Maturity Levels)** 5 级 × 5 维评估框架，作为架构工具与方法论成熟度的独立度量维度。

关键词：架构本体论；电子电气架构；软件定义汽车；具身人工智能；架构成熟度；失效哲学；Architecture Intelligence

1.2 引言：为什么“架构”是这个十年最被低估的元问题

汽车正在变成机器人。机器人正在变成移动的数据中心。数据中心正以世界模型反向重建物理世界。智能手机、可穿戴设备、智能家居正以分布式微内核统一为单一“超级终端”。

这场跨产业的“架构大迁徙”——从分布式 ECU 到中央计算、从单体 OS 到分布式微内核、从规则驱动到端到端神经网络、从工业机械臂到通用人形——并非五个独立的产业故事，而是**同一条技术曲线在五个剖面上的投影**。然而，业界对这种深层同构性的讨论仍停留在概念口号层面，缺乏系统的本体论梳理、严谨的跨域基准、以及对收敛边界的清晰刻画。

更值得关注的是：在过去十年中，“架构”这一概念被频繁地降格使用。它有时被简化为“系统框图”，有时被等同于“零部件清单”，有时被混淆为“技术选型”。这种概念稀释的代价是巨

大的——架构决策的不可逆性、跨生命周期成本、以及对组织能力的长期塑造，都在简化叙事中被遮蔽。

本文试图重新激活”架构”作为元问题的严肃性。我们从两千年的概念史出发，进入当代工程语义的严谨定义，然后以汽车 E/E 架构为锚点（因其工程约束最严苛、监管最完整、产业演进最清晰），向手机、云、机器人、具身 AI 四个邻域作横向对标，最终回到一个统一的成熟度框架。

本文不是产业咨询报告，不为任何特定企业或地理市场服务。研究问题严格限定为：**当代各产业架构在演进中表现出怎样的同构与异构？这种结构性差异由何驱动？是否存在一个统一的方法论来刻画与比较？**

1.3 第一章架构的本体论：从维特鲁威到 Architecture Intelligence

1.3.1 1.1 起源：希腊语 *archi-tekton* 与维特鲁威三原则

“架构”一词源于希腊语 *ἀρχιτέκτων* (*archi-tekton*)，由 *archi-*（首要、主宰）与 *tekton*（建造者）合成，意为“首席工匠”。公元前 30—15 年间，罗马建筑师维特鲁威（Marcus Vitruvius Pollio）在《建筑十书》（*De Architectura*）中提出了奠基性的“维特鲁威三原则”：

- **Firmitas**（坚固）——结构完整性与持久性
- **Utilitas**（适用）——功能性与目的契合
- **Venustas**（美观）——形式与体验

这一三位一体的取舍框架，在两千年后被软件架构师改写为“可靠性—功能性—可维护性”的非功能需求三角，再被系统架构师扩展为质量属性矩阵。其哲学要义历久弥新：**架构本质上是一种关于取舍的工程艺术，而非关于完美的科学。**

1.3.2 1.2 从模式语言到软件架构：20 世纪后半叶的概念跃迁

1977 年，建筑师克里斯托弗·亚历山大（Christopher Alexander）出版《建筑模式语言》（*A Pattern Language*），将“模式（pattern）”作为可复用的设计单元引入设计理论。1987 年，Kent Beck 与 Ward Cunningham 在 OOPSLA 会议上将亚历山大的模式语言移植到软件领域，1994 年 GoF 四人组（Gamma, Helm, Johnson, Vlissides）的《设计模式》编纂了 23 种经典面向对象模式，软件架构由此获得了其方法论根基。

软件架构作为独立学科兴起于 1960—70 年代的“软件危机”。关键里程碑包括：

- **1968 年北约软件工程会议**（NATO Software Engineering Conference, Garmisch）首次将“软件工程”作为术语正式确立；
- 同年，Edsger Dijkstra 发表《THE 多道程序系统的结构》，确立**分层架构（layered architecture）**范式；
- **1972 年 David Parnas** 发表《关于分解模块的标准》，引入**信息隐藏（information hiding）**原则；
- **1992 年 Perry & Wolf** 《软件架构基础》将架构研究从“写代码”提升至“系统级组织、风格、接口与约束”；
- **1996 年 Mary Shaw & David Garlan** 出版《软件架构：一门新兴学科的视角》，正式确立软件架构为独立学科；
- **2000 年 IEEE 1471** 标准发布，演化为今天的 **ISO/IEC/IEEE 42010** 架构描述标准。

1.3.3 1.3 当代严谨定义：ISO 42010、TOGAF 与 Shaw-Garlan-Booch-Kruchten 谱系

在当代工程语义中，“架构”具有高度严谨的定义。**ISO/IEC/IEEE 42010** 将架构定义为“系统在其环境中的基本概念或属性，体现在其要素、关系及其设计与演化原则中”。该标准的核心贡献在于将关注重点从“架构图”转向“架构描述 (architecture description)”，并引入了利益相关者 (stakeholders)、关注点 (concerns)、视点 (viewpoints) 与视图 (views) 的四元概念模型。

TOGAF (The Open Group Architecture Framework) 则在企业语境中将架构定义为“系统的组件结构、它们之间的相互关系，以及指导其设计和长期演进的原则与规范”。这一定义的深刻之处在于明确指出：**架构不是结构图本身，而是结构图背后的设计逻辑与治理方法。**

软件架构学界沿 Shaw-Garlan→Booch-Kruchten 的谱系逐步精炼了定义。其中，**Philippe Kruchten** 在 1995 年提出的“**4+1 视图模型**”是工程实践中最具影响力的架构表达框架：

视图	关注点	主要利益相关者
逻辑视图	功能性需求、对象/服务模型	终端用户、需求分析师
过程视图	并发、性能、可伸缩性	系统集成商
物理视图	部署拓扑、硬件分布	系统工程师、运维
开发视图	模块组织、构建依赖	开发者、项目经理
+1 场景视图	用例驱动的统一与验证	全体利益相关者

Kruchten 与 Grady Booch 进一步指出，架构涵盖了关于**系统组织的重大决策集合**：如何选择结构元素及其接口、如何定义元素协作中的系统行为、如何将基础组件复合成更复杂的子系统、以及确立指导全局组织的架构风格。这一定义揭示了架构的**不可逆性**——它定义了演化的边界与代价。

1.3.4 1.4 架构学的两条底层定律

经过半个世纪的方法论沉淀，软件与系统架构学已经收敛于两条不可动摇的底层定律：

第一定律：一切皆是权衡 (Everything is a Trade-off)

不存在“最优”的架构，只存在“在特定约束下相对合理”的架构。功能性需求与非功能性需求（可用性、可扩展性、可维护性、安全性、成本）之间始终存在结构性冲突。架构师的核心工作不是消除冲突，而是显式化冲突、量化代价、为决策提供可追溯的论证链。

第二定律：Why 凌驾于 How (Why Dominates How)

架构的战略价值不在于“如何实现”的技术选型，而在于“为何如此选择”的论证逻辑。Why 决定了系统的演化边界、容错策略与组织能力路径依赖；How 仅决定了当下一代实现的成本与性能。忽视 Why 的架构决策在长生命周期系统（汽车、机器人、操作系统平台）中代价尤其高昂。

1.3.5 1.5 架构观的当代谱系：从结构到平台到生态

经过半个世纪的演化，当代架构实践收敛于五种互补的”架构观”：

1. **结构/模块观**：强调组件分解与接口边界（继承自 Parnas 信息隐藏）；
2. **视点/利益相关者观**：强调不同角色需要不同架构视图（Kruchten 4+1）；
3. **企业治理观**：强调组织、流程、标准与投资组合（Zachman、TOGAF）；
4. **平台/生态观**：强调跨产品复用、开放接口、标准化与生态协同（Android HAL/Treble/Mainline、AUTOSAR、Kubernetes）；
5. **质量属性工程观**：强调可量化的非功能指标（ATAM、SAAM）。

1.3.6 1.6 架构演进的元规律

从历史长时段视角观察，架构演化沿着一条清晰的元规律前进：

静态集成 → 模块化 → 服务化 → 平台化 → 云边协同 → 模型驱动与 AI 原生

驱动这条曲线的并非技术潮流本身，而是复杂度上升后，组织必须把变化隔离、把能力复用、把风险显式管理。这一驱动力在汽车、机器人、具身 AI 时代尤其强烈——因为这些系统同时面临高复杂度、高安全要求、长生命周期与跨组织协作四重压力。

1.3.7 1.7 引入 Architecture Intelligence

当架构演进到 AI 原生阶段，一个深层命题浮现：**如果架构本身是一种复杂决策系统，那么是否可以用 AI 来辅助、加速、甚至自动化架构决策本身？**

我们将这一研究范式称为 **Architecture Intelligence**（架构智能，简记为 **AI²**，“**Architecture Intelligence**”的双 I 缩写）——即应用智能方法（包括知识图谱、形式化验证、大语言模型、强化学习、参考架构对标）于架构设计、约束检查、演化预测的系统性研究。需要明确区分两个相邻但本质不同的方向：

- **Architecture Intelligence (AI²)**：用 AI 做架构——研究如何让架构决策过程本身变得更智能、更可验证、更可复用；
- **Architecture for AI**：为 AI 设计架构——研究如何为 AI 工作负载（训练、推理、部署）设计最优的系统架构。

本文聚焦前者。两者虽然在工程实践中常被交织讨论，但所处理的核心问题截然不同——前者是方法论问题，后者是工程优化问题。

为什么需要 Architecture Intelligence? 当代复杂系统的架构决策已经超出了人类架构师的直觉处理能力。一辆 SDV 涉及上百个 SoC/MCU、数千个软件组件、数万条约束规则、数十种合规标准；一个人形机器人涉及数十个自由度的耦合控制、视觉-语言-动作的端到端推

理、跨场景泛化。在这种复杂度下，**架构决策的质量越来越受限于工具支持的成熟度，而不仅是架构师个体的经验。**

历史上，软件工程的每一次重大跃迁都伴随着工具的同步演化——从汇编到高级语言（编译器革命）、从瀑布到敏捷（CI/CD 革命）、从单体到微服务（容器与编排革命）。今天，复杂系统架构本身正经历类似的工具化革命：从手绘框图到模型驱动、从离散约束到形式化验证、从人工对标到 AI 辅助决策。本文在第五章将正式提出 **AI²-ML (Architecture Intelligence Maturity Levels)** 框架，作为对架构工具与方法论成熟度的系统化度量。

1.4 第二章汽车电子电气架构：从分布式 ECU 到中央计算的演进

1.4.1 2.1 概念边界：E/E 架构不只是线束与 ECU 拓扑

汽车电子电气 (Electrical/Electronic, E/E) 架构指车辆内部所有电子硬件、传感器、执行器、计算单元、通信网络与底层软件平台的总体组织方式。重要的是，E/E 架构不仅是线束和 ECU 的物理拓扑，更包括：

- **计算层**：MCU、域控制器 (DCU)、区域控制器 (ZCU)、高性能中央计算单元 (HPC)；
- **网络层**：CAN/LIN/FlexRay 向车载以太网、TSN (时间敏感网络)、SOME/IP 演化的通信骨干；
- **软件层**：RTOS/OS、AUTOSAR Classic/Adaptive、中间件、服务总线、诊断协议；
- **安全机制**：功能安全 (FuSa)、预期功能安全 (SOTIF)、网络安全 (Cybersecurity)、AI 安全；
- **运维层**：OTA 更新管理、车云协同、舰队学习闭环。

AUTOSAR 将自身定义为“标准化软件框架和开放 E/E 系统架构”。博世、Aptiv、Google、NVIDIA 等行业领军者则把软件定义汽车 (SDV) 的核心定义为：**以更少、更强的车载计算节点取代大量孤立控制器，并通过软件在全生命周期持续升级。**

1.4.2 2.2 演进的物理压力：ECU 数量危机

E/E 架构演进的直接历史背景是 ECU 数量的失控增长。博世公开数据显示，紧凑级车型今天通常配备 30—50 个 ECU，高端车型超过 100 个，部分豪华车型逼近 150 个。这一数字膨胀带来了多重危机：

- **线束复杂度**：单车线束长度可达 4—5 公里，重量 50—70 kg，成为整车 BOM 第三大成本项；
- **供应链耦合**：每个 ECU 涉及独立供应商、独立软件栈、独立验证流程；
- **验证成本爆炸**：跨 ECU 集成测试场景呈组合爆炸；
- **OTA 难以实现**：异构 ECU 的统一更新管理几乎不可能；
- **迭代周期僵化**：硬件锁定功能，软件改动需要硬件重新认证。

正是这一物理压力，推动了从“分布式功能 ECU”向“域控制”再向“区域化 + 中央计算”的平台架构演进。

1.4.3 2.3 博世六阶段演进框架

图 2.1 博世六阶段 E/E 架构演进框架。趋势为分布式 → 集中化 → 车云一体；低阶不会消失，与高阶长期共存。

博世提出并被行业广泛采纳的 E/E 架构六阶段演进框架，是当前最具影响力的产业共识：

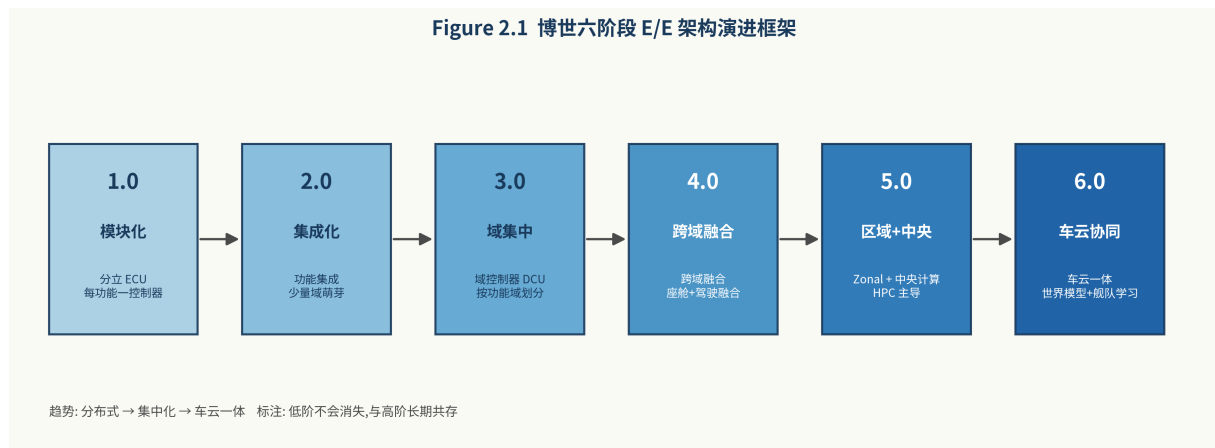


Figure 1: 图 2.1 博世六阶段 E/E 架构演进框架。趋势为分布式 → 集中化 → 车云一体；低阶不会消失, 与高阶长期共存。

阶段	名称	核心特征	典型代表
1.0	模块化 (Modular)	一项功能 = 一个 ECU; 硬软件深度耦合; OTA 几乎不可能	2010 年代前的大部分车型
2.0	集成化 (Integration)	CAN 网关 + 部分 ECU 物理逻辑集成; 初步功能复用	中低端燃油车主流配置
3.0	域集中 (Functional Domain)	5 大域 (动力、底盘、车身、座舱、ADAS); 域控制器 + 跨域网关	Audi zFAS、大众 MEB、丰田 TNGA、比亚迪 e 平台 3.0、2018—2022 年大多数 EV
4.0	跨域融合 (Cross-Domain)	舱驾融合于单 SoC; 虚拟化/Hypervisor 分区	NVIDIA Thor、高通 SA8775P、小鹏 X-EEA 3.0、理想 LEEA 3.0
5.0	中央计算 + 区域 (Central + Zonal)	1—2 个中央 HPC + 3—6 个 ZCU; 千兆/万兆以太网骨干 + TSN	Tesla HW4/AI5、Rivian R1、蔚来 Cedar、Mercedes MMA
6.0	车云协同 (Vehicle-Cloud)	算力车-云间动态分配; 端到端模型 + 世界模型; 持续舰队学习	Tesla Dojo→Cortex 2.0、华为 ADS WEWA、Waymo Foundation Model

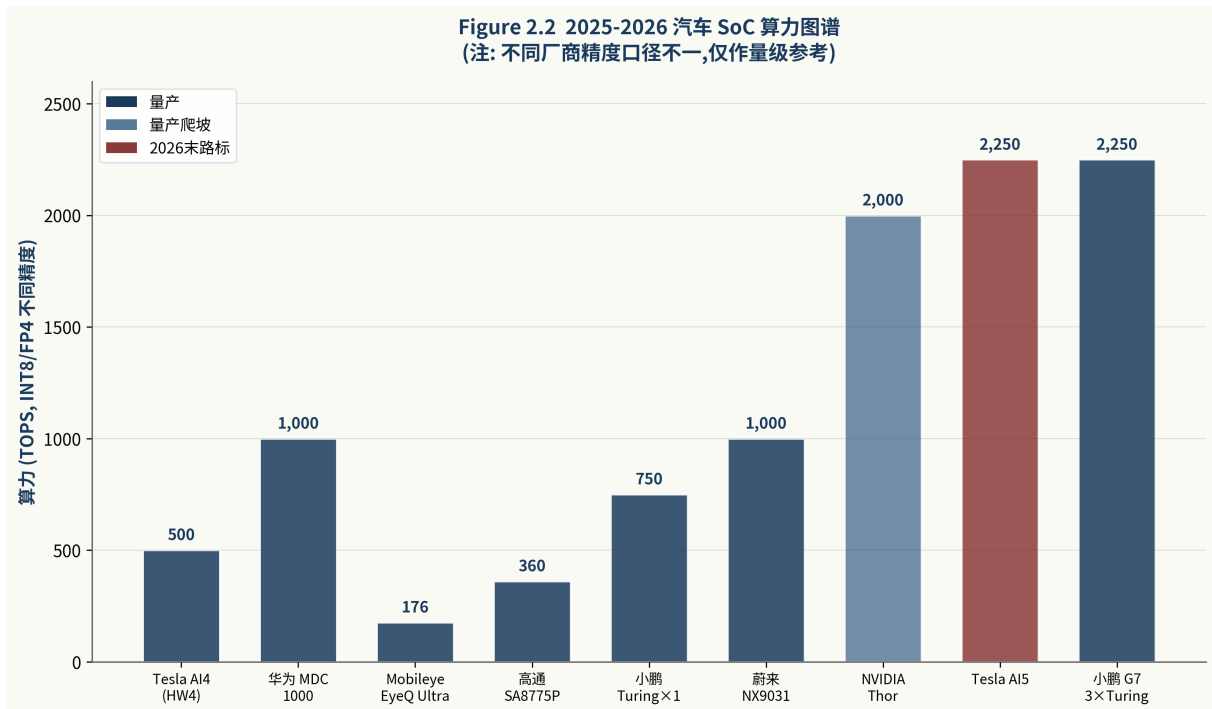


Figure 2: 图 2.2 2025-2026 汽车 SoC 算力图谱。不同厂商精度口径不一, 仅作量级参考。

麦肯锡 2023 年的研究量化了这一演进的供应链影响: **传统 ECU 市场年化收缩约 1%**, 而 **DCU/ZCU/HPC 市场以 30—40% 年增速扩张**。HPC 单价 1,000—4,000 美元, ZCU 50—70 美元。Rivian 的电气架构精简度具有标志性意义——其最新平台通过强大的中央计算单元将全车传统 ECU 从 40—80 个削减至仅 **7 个**, 并因此省去了每辆车约 **1.6 英里 (约 2.5 公里)** 的物理线束。

1.4.4 2.4 2025—2026 年主要玩家算力图谱

图 2.2 2025-2026 汽车 SoC 算力图谱。不同厂商精度口径不一, 仅作量级参考。

特斯拉 FSD 计算演进: HW3 (2019, 三星 14nm, 144 TOPS) → HW4/AI4 (2023, 约 500 TOPS) → **AI5** (2026 年末/2027 年初量产, TSMC N3P, 目标 2,000—2,500 TOPS, 首发于 Cybercab, 整颗 SoC 演化为推理 GPU, 无独立 ISP/图形单元) → **AI6** (三星 Texas 165 亿美元合同, 2025 年 7 月签约, 融合推理 + 训练单 SoC, 服务车、Optimus、Dojo 3) → **AI7** (2027—28, 定位”太空 AI 计算”)。九个月的芯片迭代节奏在产业内独一无二。

NVIDIA DRIVE Thor: 2024—2026 年量产爬坡, 2,000 TFLOPS FP8, 77 亿晶体管, Arm Neoverse V3AE “Poseidon” CPU + Blackwell GPU + Transformer Engine; 通过 MIG (多实例 GPU) 实现座舱与 ADAS 在同芯片的域隔离; ASIL-D / ISO 26262 / ISO 21434 全栈认证。客户覆盖比亚迪、小鹏、理想、Zeekr、奔驰、JLR、小米 (YU7 与 2026 款 SU7 全系搭载 Thor-U 700 TOPS 单芯版本)。

高通 Snapdragon Ride Flex: SA8775P (5nm, 舱驾融合, ASIL-D) + SA8650P (ADAS 专用) + SA8295P (座舱平台, 已大规模搭载于奔驰、蔚来、小鹏、Zeekr、小米、

Leapmotor)。

Mobileye EyeQ6 High (2024 年末量产, 2025 年部署) 与 **EyeQ Ultra** (12 个双线程 RISC-V 核 + 256 GFLOPS Arm GPU + 64 个专有加速核, <100W, 单芯片 L4 目标 2025—26)。Mobileye 的差异化战略是 REM 众包高精地图、RSS 形式化安全模型、True Redundancy (视觉与雷达独立子系统冗余)。

中国新势力自研芯片的三大突破:

- **蔚来神玑 NX9031** (2024 年 7 月 NIO IN 发布): 业界首颗 5nm 车规自动驾驶芯片, 50+ 亿晶体管, 32 核 big.LITTLE CPU, LPDDR5x 内存带宽 **546 GB/s**, 相当于 4 颗 NVIDIA Orin X, ASIL-D 安全岛, 毫秒级双芯故障切换。
- **小鹏图灵芯片** (2024 年 8 月流片, 2025 年 Q2 量产): 40 核处理器 + 2 颗自研 NPU, 单芯 **750 TOPS**, 支持 30B 参数本地大模型; 小鹏 G7 Ultra 搭载 3 颗 Turing = **2,250 TOPS**。
- **华为 MDC 1000** (基于昇腾 910B): 1,000 TOPS, 支持 16 摄像头 + 3 激光雷达, 双 MDC 1000 冗余 <50ms 切换, ASIL-D。

端到端神经网络的代际跃迁:

- **理想 VLA 模型** (2025 年 7 月 i8 首发): 首个量产车端 Vision-Language-Action 模型, **12 亿公里真实车数据训练**, Mind GPT-3o 多模态端到端 3 万亿 token 训练。
- **华为 ADS 4.0** (2025 年 4 月 22 日发布, **WEWA 架构**): World Engine 云端生成 1,000 倍极端场景密度 + World Action Model 车端原生驾驶基础模型; 端到端时延 -50%、通行效率 +20%、急刹 -30%。截至 2026 年 1 月累计 87.6 亿公里辅助驾驶里程。
- **比亚迪天神之眼 / DiPilot** (2025 年 2 月, 全系免费): DiPilot 100/300/600 三档分级; 日均云端训练数据 7,200 万公里, 来自 44 万 +L2/L3 车辆舰队。

1.4.5 2.5 大众 CARIAD 困局与 Rivian 联盟: 垂直闭环受挫到平台联盟的转型范例

如果说 Tesla 代表了垂直整合的成功范例, 那么大众集团 CARIAD 项目则是垂直闭环战略受挫、被迫转向平台联盟的典型案列。理解这一案例对于把握汽车产业架构演进的复杂性至关重要。

E3 架构的代际困境: CARIAD 致力于打造横跨大众集团所有品牌的统一软件堆栈 (Group Software Stack), 整合四个核心技术领域: 集团驾驶栈 (自动驾驶)、体验栈 (座舱体验)、云栈 (车联网生态)、运动栈 (车辆动态控制)。其架构演进规划为:

- **E3 1.1:** MEB 平台支撑, 部分 HCP (高性能计算平台) 集成;
- **E3 1.2:** PPE 平台 (奥迪 Q6 e-tron、保时捷 Macan EV) 应用, 5 个 HCP 覆盖驾驶、辅助、信息娱乐与安全;

- **E3 2.0**: 原计划专为 SSP (可扩展系统平台) 研发, 目标是真正意义上的中央计算 + 区域架构。

然而 E3 2.0 的研发严重滞后, 迫使大众做出三项重大架构决策:

决策一: 地缘分裂战略——SDV West 与 SDV East。在西方市场, 大众与 Rivian 于 2024 年 11 月成立 50:50 合资公司”RV Tech”, 大众注资 58 亿美元, 全盘引入 Rivian 的中央区域硬件架构与软件技术栈。首款应用车型 ID.EVERY1 计划 2027 年量产。在中国市场, 大众选择与小鹏汽车深度合作, 共同开发 CEA (China Electronic Architecture) 以适配中国本土数字化生态与极度内卷的迭代节奏。

决策二: Rivian 架构的颠覆性精简。Rivian R1 / R2 的电气架构通过强中央计算将传统 40—80 个 ECU 削减至仅 7 个, 省去约 2.5 公里物理线束。借助虚拟化技术整合 ECU 可使单车 BOM 成本降低 800—1,500 美元。

决策三: 与高通战略联盟。2026 年 1 月, 大众与高通签署意向书, 将 Snapdragon 平台引入下一代座舱与 ADAS 系统。

这一案例的架构启示极具代表性: 当垂直闭环的内部研发能力不足以支撑代际架构跃迁时, 平台联盟成为唯一可行的路径。但平台联盟的代价是部分丧失对架构控制点的掌控权——Rivian 拥有底层电气架构与软件栈、高通拥有座舱与 ADAS SoC、大众则收缩到”整车产品定义 + 品牌 + 制造”的边缘角色。这正是 SDV 时代价值链重排的微观映射。

1.4.6 2.6 参考框架与标准矩阵

软件框架双轨: AUTOSAR Classic(信号驱动、深嵌入式 ECU、OSEK RTOS)与 **AUTOSAR Adaptive** (SOA、ARA::COM SOME/IP、POSIX, 面向 SDV/AD/HPC) 当前共存于 R24-11 版本。前者面向硬实时与功能安全约束, 后者引入服务化与高算力。

云原生汽车栈: SOAFEE (Arm 主导, 2021 成立, 博世/大陆/采埃孚/AWS/Red Hat 加入) 将云原生范式 (容器、Kubernetes 编排) 引入混合关键性汽车边缘。**COVESA** (原 GENIVI, 2022 改名) 维护 VSS (Vehicle Signal Specification) 与 vSomeIP 开源实现。

机器人栈进入车辆: ROS 2 Jazzy Jalisco (2024 年 5 月, 当前 LTS) 通过 Apex.AI (ASIL-D 认证 ROS 2) 和 Autoware 进入生产车辆。2024—2026 年学术界与产业界已就 **AUTOSAR AP** 与 **ROS 2** 互补形成共识, 主流方案是通过 SOME/IP-DDS 网关桥接。

安全与合规矩阵:

标准	范围	状态
ISO 26262 (2018 v2)	功能安全; ASIL A—D	基础标准
ISO 21448 SOTIF	预期功能安全	ADAS/AV 必需
ISO/SAE 21434 (2021)	网络安全工程全生命周期	UN R155 法定底标
UN R155	网络安全管理体系 (CSMS)	2024 年 7 月对所有新车强制
UN R156	软件更新管理体系 (SUMS)	同上

标准	范围	状态
ISO 24089 (2023)	软件更新工程	R156 工程化
ISO/PAS 8800 (2024)	汽车 AI 安全	2024 年 12 月发布
欧盟 AI Act (2024/1689)	高风险 AI 系统	车辆 AI 安全部件 2027 年全面适用
欧盟 Data Act (2023/2854)	互联设备数据访问	2025 年 9 月适用
欧盟 CRA (2024/2847)	数字产品韧性	2027 年 12 月全面适用
GB 44495/44496/44497-2024 (中国)	车辆网络安全/软件更新/数据存储	2026 年 1 月新车型强制

这一标准矩阵的深层含义是：**汽车架构的演进，不是向手机单纯靠拢，而是向“可升级但可审计、可智能但可证明”的受监管平台靠拢。**这是汽车与消费电子在架构哲学上保持长期分化的根本原因。

1.5 第三章跨产业架构对标：五维矩阵、失效哲学与两层结构

1.5.1 3.1 五维基准矩阵

下表浓缩了五大产业在硬件、软件、数据/AI、生态、安全五个维度上的架构选择：

维度	汽车 E/E	智能手机/消费电子	互联网/云	机器人	具身 AI
硬件	Thor 2,000 TOPS、HW5/AI5、MDC 1000、舱驾融合 SoC + 区域控制器	A19/M5 神经引擎 45 TOPS、骁龙 8 Elite Gen5、Kirin 9030	Blackwell GB200 NVL72 1.44 EFLOPS FP4、TPU v7 Ironwood 4.6 PFLOPS、Trainium2	Jetson Thor 2,070 TFLOPS FP4、电驱/腱驱执行器、相机 +LiDAR+ 触觉	同上 + 云训练 (Cortex/Dojo 3、CloudMatrix 384)
软件	AUTOSAR Classic/Adaptive、QNX、Linux、Tesla 自研、HarmonyOS-Auto	iOS/Darwin、Android/Linux、HarmonyOS NEXT 微内核、HyperOS 混合内核	Linux + Kubernetes (2025 年生产采用 82%) + 服务网格 + 微服务 + Serverless	ROS 2 DDS + Nav2 + MoveIt2、AUTOSAR-AP 桥接	VLA + System1/System2、Diffusion Policy、Flow Matching
数据/AI	车端端到端 NN (FSD V14、ADS 4.0 WEWA、Mind GPT VLA) + 云训练舰队学习	端侧 3—8B 模型 + 私有云算 (Apple PCC、AICore、Pangu) + LoRA	千亿—万亿参数训练，FP8/FP4 精度	GR00T N1、π0/π0.5、Helix、Gemini Robotics、Skild Brain、OpenVLA	NVIDIA Cosmos 世界模型 + Isaac Sim + Omniverse + Wayve GAIA-3

维度	汽车 E/E	智能手机/消费电子	互联网/云	机器人	具身 AI
生态	OEM 全栈 (Tesla) / Tier-1 联盟 (华为 HIMA) / 平台 (NVIDIA、高通)	围墙花园 (Apple、华为) / 平台共享 (Android)	开源主导 (CNCF 15.6M 开发者) + 超大规模商	垂直整合 vs 水平平台	NVIDIA 三计算机范式 (训练-仿真-运行)
安全	ISO 26262 + 21434 + R155/156 + ISO/PAS 8800 强制	TEE/Secure Enclave + GDPR/PIPL + E2E 加密; 无 FuSa	SOC 2 / ISO 27001 / 零信任 / 机密计算	ISO 13482/10218 (2025 更新) /13849	标准空白; DeepMind ASIMOV Benchmark 等新提议

1.5.2 3.2 硬件趋同：全球架构基底的收敛

硅基平台呈现明显的跨产业收敛趋势。**Arm Neoverse V3AE** 同时出现在 **NVIDIA Drive Thor、Grace、Vera** 与 **AWS Graviton4** 中；**Blackwell GPU IP** 同时支撑 **B200 数据中心 GPU** 和 **Drive Thor 车载 SoC**。这是历史上第一次出现汽车与超算共享同一架构基底——这意味着汽车 OEM 与云超大规模商之间的能力溢出将发生**双向流动**。

异构计算 (CPU + GPU + NPU + DSP + 专用加速器) 已经成为所有产业的默认范式。差异不在于“是否异构”，而在于**算力配比、确定性保证、热设计与成本目标**。

1.5.3 3.3 失效哲学：fail-soft vs fail-operational 的根本分水岭

硬件趋同，但软件分化。真正的差异不在“是否有 OTA”或“是否用 Linux/Android”，而在**失效哲学**。

手机和互联网更接近 fail-soft + 快速恢复：服务崩溃可以毫秒级切换至备用节点；应用闪退可以立即重启；OTA 失败可以回滚。这种“允许失败并通过冗余兜底”的软实时架构在虚拟世界中无可挑剔。Android 可以用 Mainline 把系统模块持续更新，Kubernetes 可以滚动升级和回滚。失效的代价是用户体验下降，而非物理损害。

汽车和很多机器人则要求 fail-safe / fail-operational + 可解释责任归属：当传感器检测到致命障碍物时，计算平台向制动卡钳下达指令的时间窗口是绝对封闭的；当一个执行器失效时，系统必须降级到安全状态而非崩溃；当事故发生时，必须有可审计的决策链来归责。汽车更新必须经过安全分析、版本匹配、法规与供应链联动，且**更新后的行为不允许破坏已验证的安全边界**。

这一失效哲学的差异，是汽车与消费电子在架构层面长期分化的**真正根源**——比实时性更深、比合规更早。它直接决定了：

- 为什么汽车不能单纯运行 Linux，而必须在底层嵌入 AUTOSAR 或微内核 RTOS；
- 为什么 OTA 在车端比手机端复杂十倍；
- 为什么 ASPICE、ISO 26262、ISO 21434 不可能被消费电子流程替代；
- 为什么端到端神经网络在车端必须配以可解释中间层与运行时监视器。

1.5.4 3.4 Hypervisor 作为数字-物理鸿沟的桥梁

汽车与高自由度机器人的架构系统面临一个深刻的矛盾：它们一方面**强烈渴望拥有 Android/Linux 级别的开源繁荣应用生态**，以承载需要海量内存的 AI 大模型；但另一方面，又在生死关头**绝望地依赖 AUTOSAR/RTOS 级别的硬实时干预机制**，以确保即便在内存溢出或上层应用死机时，制动卡钳与机械臂依然能够执行绝对安全的物理指令。

在这场“鱼与熊掌”的较量中，**Hypervisor 虚拟化层成为当前唯一可行的通用解药**。底层运行严格遵守 ISO 26262 ASIL-D 的 AUTOSAR Classic，专注硬实时的制动、转向控制；上层运行 Android Automotive、Linux 或定制车机 OS，专攻座舱娱乐与导航；二者在同一个强大的 SoC 上被 Hypervisor 严格隔离，形成“互通又互不干扰”的混合关键性架构。

QNX Hypervisor、COQOS、Hypervisor for ARM TrustZone、以及开源的 Xen Project Automotive 是当前主流选择。NVIDIA Drive Thor 通过 MIG (Multi-Instance GPU) 将这一虚拟化范式从 CPU 域扩展到 GPU 域，使得 ADAS 与座舱在同一颗 Blackwell GPU 上获得硬件级隔离。这是汽车架构史上具有标志意义的技术突破。

1.5.5 3.5 数据/AI 架构的边-云分工：双向收敛

智能手机（端侧 3—8B 参数 + 私有云计算）与汽车（车端 VLA + 云端世界模型）在**架构模式上几乎相同**：受限设备做实时推理，云做训练 + 仿真 + 难例数据生成。

NVIDIA Cosmos 世界模型已被 Toyota (DriveOS)、1X、Skild AI、Figure AI、Agility、Uber、Waabi、Foretellix、小鹏、Galbot、Fourier 同时采用——汽车与机器人正在共享同一套世界模型基础设施。这是这场架构革命中最被低估的事实。

具体的边-云分工策略呈现三种典型范式：

1. **小模型本地 + 大模型云端** (Apple Intelligence、Google Tensor + Cloud TPU)：3—8B 参数模型驻留端侧，触发性地调用云端 Gemini/GPT 级大模型；
2. **完整模型本地推理 + 增量学习云端** (Tesla FSD、华为 ADS)：端侧运行完整端到端 NN，云端仅做训练与版本下发；
3. **端云双模型协同** (Waymo System 1/System 2，详见第四章)：端侧快思考模块 + 云端慢思考 VLM，按场景复杂度动态切换。



Figure 3: 图 3.2 双层趋同-分化模型。第一层共享基础设施持续趋同; 第二层行业特定上层持续分化。

1.5.6 3.6 生态: 三种集结哲学

跨产业观察可识别出三种生态构建哲学:

- **垂直收敛型** (Tesla 范式): 从硅、OS、应用到云、机器人单一企业全栈自研;
- **水平微内核 + 大模型平台型** (华为范式): 一个微内核 OS + 一个基础大模型 + 一个 AI 加速器底座, 覆盖手机—PC—车—云—具身机器人;
- **联邦平台 + 最佳合作伙伴型** (Google 范式): TPU + Gemini 提供给内部产品与外部合作伙伴。

每种哲学背后是不同的”架构控制点”持有策略。我们在第六章将系统化这一概念。

1.5.7 3.7 两层结构: 基础设施趋同, 控制语义与证明义务分化

图 3.2 双层趋同-分化模型。第一层共享基础设施持续趋同; 第二层行业特定上层持续分化。

综合五个维度的对标, 本研究的核心判断是: **未来不会完全”融合成一种架构”, 而会形成两层结构。**

第一层: 共用底座 (趋同)

- 异构算力 (CPU + GPU + NPU + 专用加速器)
- 以太网骨干 + TSN
- 虚拟化与混合关键性 OS
- 服务化接口 (SOA、SOME/IP、DDS、gRPC)
- 数字孪生与世界模型仿真

- 端到端模型训练与部署管线
- OTA 流水线与持续部署

这部分将明显收敛，跨产业边界将变得模糊。SoC 厂商、云平台、开源生态将主导这一层。

第二层：行业特定上层（分化）

- 车辆/机器人物理控制层
- 安全壳与失效模式管理
- HMI 与用户交互范式
- 运维策略与生命周期管理
- 责任证明与可审计性
- 法规适配与合规框架

这部分将继续分化，并因监管强化而进一步固化。Tier-1、OEM、机器人本体厂商、领域专家将主导这一层。

用一句话概括：基础设施趋同，控制语义和证明义务分化。

这一双层结构判断比简单的”会收敛/会分化”二元论更精确。它解释了为什么我们既看到 Tesla 把 FSD 计算栈无缝迁移到 Optimus（共用底座），又看到汽车与机器人在功能安全、责任框架、HMI 上保持独立演化路径（行业特定上层）。

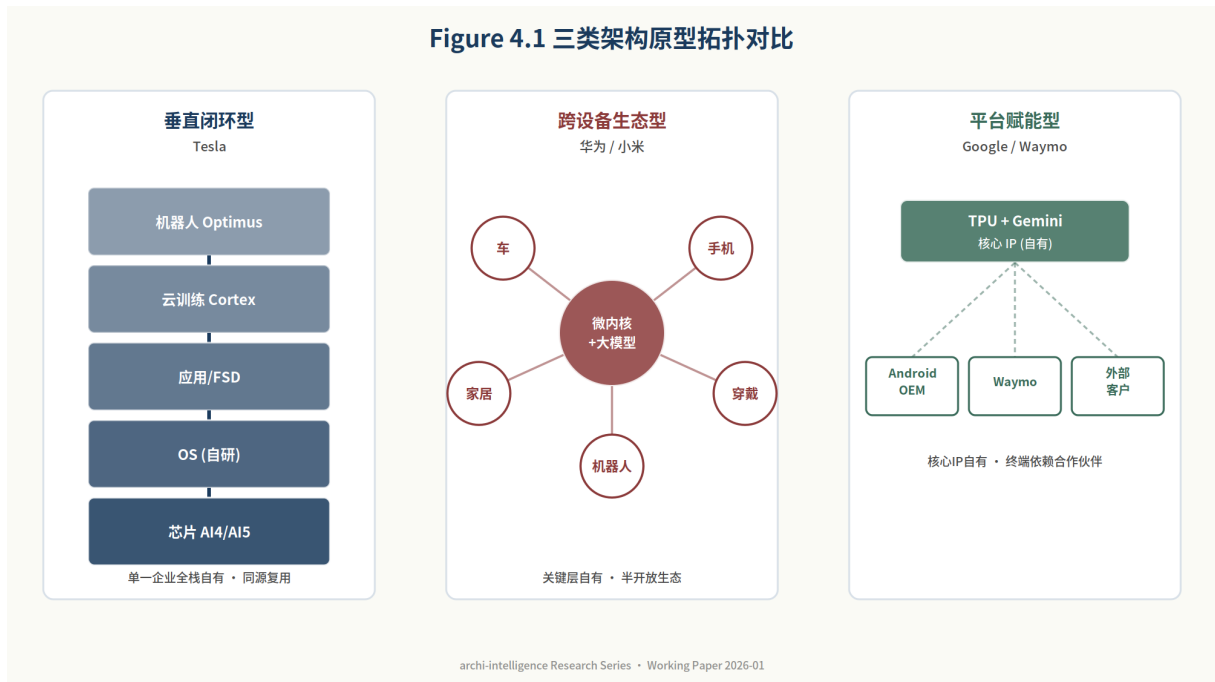


Figure 4: 图 4.1 三类架构原型拓扑对比。垂直闭环型 (Tesla)、跨设备生态型 (华为/小米)、平台赋能型 (Google/Waymo) 三种形态。

1.6 第四章三类架构原型的深度基准

基于第三章的生态分类，本章对三类架构原型进行深度案例剖析。每个案例围绕”架构哲学—复用机理—差异点—限制条件”四个层次展开。

1.6.1 4.1 垂直闭环型：Tesla 的形态无关感知栈

图 4.1 三类架构原型拓扑对比。垂直闭环型 (Tesla)、跨设备生态型 (华为/小米)、平台赋能型 (Google/Waymo) 三种形态。

架构哲学：一颗 AI 芯片 + 一个端到端神经网络 + 一套形态无关感知，跨越车、人形、Robotaxi、能源。

4.1.1 硬件复用：从 FSD 到 Optimus 的同源芯片 Optimus 的计算大脑 (Bot Brain) 直接复用了成熟的汽车 FSD 硬件生态。当前主要依赖低成本但极其高效的 AI4 推理芯片，未来将迁移至 AI5/AI6。这种复用并非”凑合可用”，而是 Tesla 在 FSD 芯片设计时锱铢必较地优化了时序、功耗与每瓦特硅性能 (performance-per-watt)，这种极致的功耗比恰好契合了受限于电池容量的 Optimus 机器人 (其躯干内置 2.3 kWh 电池组，目标 8 小时续航)。

通过底层流水线计算调度与高效内存访问，高频传感器数据得以在极低延迟下跨进程共享。AI5 芯片设计时已经显式考虑了机器人推理负载，使得同一颗 SoC 可以服务车与人形两种形态。

4.1.2 视觉感知与神经网络的直接迁移 Tesla 坚持**纯视觉路线**——汽车与机器人均无激光雷达。这一备受争议的架构决策是 FSD → Optimus 复用的**核心技术前提**：

- **8 摄像头视觉感知配置**：Optimus 几乎完全沿用同样的相机组合；
- **占用网络 (Occupancy Networks, 2022 AI Day 公开)**：体素化 3D 稠密占用 + 占用流场，从相机生成形态无关的 3D 场景表示，**体素大小靠近机器人时可降至 10 cm**——这是 FSD 栈可以迁移到 Optimus 的核心技术原因；
- **48 个独立神经网络**：FSD 完整构建包含 48 个独立 NN，消耗约 70,000 GPU 小时训练，每个时间步长输出 1,000 个离散预测张量；
- **FSD V12 (2024 年 1 月) 首个端到端 NN**，替代 30 万 + 行 C++ 控制代码；**V13 (2024 年 12 月) 原生 HW4 分辨率**，数据量 4.2×；**V14 (2025–26) 参数 4.5× V13**。

这套在二维道路上用于语义分割、目标检测、单目深度估计、鸟瞰图、占用网络的顶尖视觉算法，被几乎原封不动地赋予了 Optimus。使得机器人无需重新发明轮子，便直接获得了世界级的三维环境感知与物体识别能力。

4.1.3 运动学维度的不可复用：差异的工程颗粒度 尽管拥有相同的”大脑”和”眼睛”，但汽车与人形机器人在”小脑”（底层运动控制）与执行器硬件上存在云泥之别。

运动学维度的本质差异：自动驾驶的轨迹规划是在低维空间内进行的（车辆在道路平面的轨迹规划），运动学模型相对简单；而 Optimus Gen 3 拥有 **22 个手部自由度 + 3 个腕部自由度（每臂 25 DoF, 接近人手 27 DoF） + 全身 28 个自由度**，其行走步态涉及重心动态转移、双足力反馈平衡、防摔倒阻抗控制——这是一套完全独立于驾驶逻辑的全新物理控制体系。

机械致动器的激进设计：Optimus 装备了六种 **Tesla 全自研定制致动器**（三种旋转减速器 + 三种线性致动器），通过无刷电机与行星滚柱丝杠（Planetary Roller Screw）结合，单个致动器在仅 2 英寸短行程内能爆发出举起 500 kg 三角钢琴的推力。**Gen 3 手部 50 个执行器、腱驱 (tendon-driven) 所有执行器置于前臂，触觉指尖灵敏度 4× Gen 2**。在专利 WO2024/073138A1 中，Tesla 为机械手设计了精妙的欠驱动线缆结构——仅用 6 个致动器（拇指独占 2 个、其余四指各 1 个）配合前置线缆与扭转弹簧，便实现了对 11 个手部关节的高效精密控制，能够完成捏拿鸡蛋的精细动作。

4.1.4 训练基础设施的同源复用 Tesla 为 FSD 打造的大规模数据标注管道、硬件在环仿真系统、舰队级评估引擎，全部被复用于机器人闭环训练。这种基础设施层面的复用极大加速了实用化进程。

训练算力演进：Dojo D1 (2021, 7nm, 362 TFLOPS BF16/CFP8) → D2 (2025 年 4 月 TSMC 晶圆级量产) → **2025 年 8 月 Dojo 项目解散** (Peter Bannon 离职, 约 20 工程师创立 DensityAI) → **Cortex 集群** (Austin, 约 6.7 万 H100 等价算力) → **Cortex**

2.0 (Giga Texas, 2026 年 4 月 250 MW 启动, 2026 年中 500 MW 满产) → **Dojo 3** 自 2026 年 1 月以 AI5/AI6/AI7 SoC 板形式重启。

4.1.5 组织层面的复用：FSD 与 Optimus 团队合并 2025 年 6 月, Optimus 项目领导权由 Milan Kovac 转交 Ashok Elluswamy——后者同时担任 FSD/Autopilot 副总裁。这是 FSD 与 Optimus 团队事实合并的官方信号, 也是垂直闭环型架构在组织层面的极致表达。

4.1.6 限制条件 Tesla 范式的核心限制是:

1. **极高的资本与人才门槛**——同时垂直整合硅片、OS、应用、云、机器人需要数百亿美元的投入;
2. **纯视觉路线的安全争议**——监管机构对仅相机配置的 L4 责任声明仍持保留态度;
3. **责任模型的脆弱性**——一旦端到端 NN 出现难以解释的故障模式, 垂直整合反而使追责更困难;
4. **生态封闭的代价**——缺乏第三方开发者生态, 应用创新受限于 Tesla 单一研发节奏。

1.6.2 4.2 跨设备生态型：华为与小米的两条中国路径

跨设备生态型的核心特征是从拥有海量用户的智能手机和智能家居切入, 将千锤百炼的底层操作系统和算法池化, 进而以极低的边际成本”降维赋能”汽车与具身机器人领域。华为与小米代表了这一原型在中国的两条平行路径。

4.2.1 华为：制裁倒逼的水平微内核 + 大模型平台 架构哲学：一个鸿蒙微内核 + 一个盘古大模型 + 一个昇腾算力底座, 覆盖手机—PC—车—云—具身机器人。

硬件：Kirin 9000s (SMIC N+2, 2023) → 9020 (2024) → **9030** (SMIC N+3, 中国首个 5nm 级量产硅, 2025 年 TechInsights 确认); Ascend 910B (376 TFLOPS BF16) → 910C (双 die 780 TFLOPS); MDC 610/810/1000; **CloudMatrix 384 超节点**——384 颗 Ascend 910C 通过 6,912 个 400G OSFP SiPh LPO 光模块全互联, 约 300 PFLOPS BF16 (1.67× NVIDIA GB200 NVL72), HBM 49.2 TB (3.6×), 功耗 559 kW (3.9×)。“以光互联 + 集群规模换单芯代差”是华为面对算力封锁的核心战法。

软件：HongMeng 微内核 + Ark Compiler/Runtime + ArkTS + Cangjie 仓颉语言 + BiSheng NDK。HarmonyOS NEXT (5) 于 2024 年 10 月 22 日商用, 完全剥离 AOSP 与 Linux 内核; HarmonyOS 6 于 2025 年 11 月 25 日随 Mate 80 发布。截至 2025 年 11 月, HarmonyOS 5/6 装机超 2,700 万台, 原生应用与元服务 30 万 +, 注册开发者 1,000 万 +。其核心架构武器是分布式任务调度系统与超级终端 (Super Device) 抽象层, 使手机、手表、智慧屏、车机能够实时共享彼此的算力、摄像头与传感器数据。

数据/AI：Pangu 5.0 → 5.5 (718B 参数深思模型 + 五个垂直行业子模型, 能够在 5 分钟内完成超过 10 步的复杂逻辑推演); Pangu 世界模型生成 4D BEV 视频 + LiDAR 点

云用于驾驶与具身 AI 训练。

Kuafu (夸父) 人形机器人: 由乐聚机器人公司硬件支持, 2024 年华为开发者大会 (HDC) 首次展示。架构精妙之处在于”云边端”极致耦合:

- 底层运行 OpenHarmony 操作系统;
- “灵魂”完全接入华为云端的 **Pangu Embodied Intelligent Large Model**;
- Pangu 强大的自然语言理解、任务规划与生成能力被复用来指导双臂协同任务 (扫地、做饭);
- Pangu 自动生成操作教程与训练视频供机器人快速学习, **大幅缩短技能泛化周期**;
- **6G 网络低延迟特性**被规划融入下一代机器人架构, 以突破单机算力的物理瓶颈。

全栈生态: 三种商业模式——纯零部件 (Kirin/Ascend/MDC) + HI (华为 Inside, Avatr、奥迪 Q6L e-tron) + **HIMA 鸿蒙智行** (问界/智界/享界/尊界)。**乾崮智驾 2025 年 8 月装车破百万, 2025 年底 140 万辆, 2026 年目标 300 万辆/80+ 车型。**

4.2.2 小米: 消费品牌驱动的垂直整合 OEM 架构哲学: 一个 HyperOS 双内核 + 一个 HyperConnect 协议 + 一组全球最佳供应链硅片, 构建”人 × 车 × 家”生态。

硬件: 手机用骁龙 8 Elite Gen 5 旗舰 + 2025 年自研 **Xring O1** (3nm 首款应用处理器); 汽车用 **NVIDIA Drive Thor-U 700 TOPS** 单芯 (2026 款 SU7 与 YU7 全系标配) + 高通 8295P 座舱 SoC; 激光雷达禾赛 AT128 + 4D 毫米波 + 11 摄像头 + 12 US\$。

软件——双内核异构基座: HyperOS 1.0 (2023 年 10 月) → 2.0 (2024) → 3.0 (2025—26)。底层采用**双内核架构**——Linux (手机/平板/车) + **Vela RTOS** (小米自研, 基于 NuttX 衍生, 自 2017 年起赋能 IoT, 支持 200+ 处理器/20+ 文件系统)。框架层融合 AOSP + Vela + 8 个新子系统 (含 AI 子系统)。**HyperConnect 跨端协议与 HyperOS Cabin** (Android 12 + Linux 5.4 内核) 支撑跨设备协同。

CyberDog 与 CyberOne 的复用红利:

- **CyberDog (仿生四足, 2021; 2.0 在 2023)**: 搭载 NVIDIA Jetson Xavier NX 边缘超级计算机 (含 **384 个 CUDA 核心 + 48 个 Tensor 核心**), 处理 11 个高精度传感器数据;
- **视觉算法直接迁移**: 通过 HyperOS 统一接口, 小米手机部门积淀深厚的成像技术 (AI 交互相机系统、超广角鱼镜头畸变校正、与徕卡合作沉淀的光学算法) 直接赋能机器人, 使其获得顶级环境感知与深度建模能力;
- **CyberOne (全尺寸双足人形, 2022 年 8 月)**: 177 cm / 52 kg / 21 DoF, **峰值关节扭矩高达 300 Nm**。为支持复杂双足平衡, 小米在 HyperOS 之上专门为其开发了**基于阻抗控制的机械动力学解算引擎**——这正是手机”看懂”与机器人”看懂并产生精确物理反作用力”的控制维度差异所在。

数据/AI: MiLM 大模型族 + MiMo 7B 推理模型 (2025 开源); XLA 认知大语言模型 集成于 2026 款 SU7; Xiaomi HAD 端到端 NN, 2024 年覆盖 100 城 NOA, 2026 全系标配。

全栈生态: 垂直整合 OEM—9,100 吨 HyperCasting 一体压铸 + 自研钛合金 Titans Metal + 自研 V6/V6s/V6s Plus 电机 + 联合工程 800V 电池 + 直营门店。SU7 上市 22 个月累计 60 万 + 辆; 2026 年 EV 目标 55 万辆; SU7 在 2025 年中国销量 25.8 万辆, 超越 Model 3 的 20.0 万辆。

4.2.3 为何架构选择截然不同?

维度	华为	小米	根本原因
内核	纯微内核 (HongMeng)	混合 (Linux + Vela + AOSP)	制裁倒逼 vs 上市速度优先
商业模式	Tier-1 (HIMA)	整车 OEM	任正非政策禁止造车 vs 资本充裕 + 消费品牌渠道
硅片策略	全自研 (HiSilicon) + SMIC	全球最佳采购 + 选择性自研	制裁封锁 vs 时间-成本最优
模型路径	Pangu 自研 + 自有训练算力	MiLM 自研 + 选择性外采	AI 基础设施战略主权 vs 应用优先

架构复用点: 华为是行业内最干净的**横向跨域复用典范**——同一 HongMeng 微内核与同一 Pangu LLM 在手机、PC、车机、ADS、Kuafu 机器人下都运行。ADS WEWA = Pangu 世界模型驾驶专业化版; Kuafu = Pangu 具身专业化版。小米则采用**混合复用**: HyperOS + HyperConnect 共享, 但底层异构 (手机骁龙、车机 NVIDIA、机器人 NVIDIA Jetson)。

1.6.3 4.3 平台赋能型: Google 的联邦 AI 平台与 Waymo 的可证明安全架构

平台赋能型的核心特征是**通过开放平台、芯片、工具链、模型对外赋能, 自身不必拥有所有终端形态**。Google 是这一原型的极致代表。

4.3.1 Google 的架构哲学: TPU + Gemini + 垂直合作伙伴 硬件: Pixel 10 (Tensor G5, TSMC 3nm, 2025 年 8 月, 战略放弃三星代工); **TPU v7 Ironwood** (2025 年 11 月 GA, 4,614 TFLOPS FP8 单芯, 略超 NVIDIA B200, 192 GB HBM3e, 9,216 芯 superpod = 42.5 EFLOPS FP8, 30x perf/W vs 初代 TPU, **Gemini 3 完全在 TPU 上训练**——脱离 NVIDIA 依赖的关键节点); Axion ARM CPU (基于 Neoverse V2)。

生态合作: Anthropic 签下 100 万颗 Ironwood 大单; Apptроник 在 2026 年 2 月 Series A 扩展轮中获得 Google + B Capital 领投的 5.2 亿美元 (估值 50 亿美元) ——Apollo 人形机器人搭载 Gemini Robotics + Gemini 3, 客户包括奔驰、GXO、Jabil、John Deere。

4.3.2 AAOS 的架构重构: 从 C++ 地下室到 TaskView 层级架构 Android Automotive OS (AAOS) 经历了一次具有标志性意义的内部架构重构, 深刻揭示了”屏幕主权”在 SDV 时代的战略价值。

旧架构: C++ 地下室 (Sinkhole) 模式。在 Android 11/12 的 Native Era, 投屏系统采用极度僵化的设计: 手机端发送视频流, 车机通过 C++ 原生服务在硬件抽象层 (HAL) 初始化图层表面 (Surface), 在系统 UI 上粗暴地”打个洞 (punch a hole)”, 直接将视频缓冲管线对接至底层显示硬件控制器。这种架构的致命缺陷在于——**Android 窗口管理器 (WindowManager)** 对屏幕上的这块投屏区域彻底失去了感知。如果车辆突发电池告警或主动刹车提示, 系统无法在屏幕顶层安全地叠加通知, 引发 UI 冲突与黑屏死机隐患。

新架构: Layered Architecture + Concurrent Multi-User。从 AAOS 14 开始, 旧的底层 C++ 视频流管道被彻底废弃, 取而代之的是分层架构与并发多用户框架。投屏应用被提升至 Java/Kotlin 应用层。当手机连接时, **CarProjectionManager** 充当调度员, 在车机内自旋启动一个名为 **TaskView** 的虚拟容器, 将投屏视作原生的 Android 独立任务运行其中。

架构重构的深层战略意义:

1. **OEM 重新夺回屏幕”每一颗像素”的完全主权**——能够自由将投屏窗口限制在特定区域, 并在其上方叠加原生空调控制或安全警报 UI;
2. **配合 DisplayManager 机制**, 乘客可以在多个车载屏幕之间随意流转虚机任务, 打通多终端生态协同;
3. **位于手机和车辆硬件之间的标准框架层**, 为 OEM 提供了坚固的检查站, 用以严格”把关”车辆私密传感器数据向手机生态倒灌的权限。

这一案例揭示了平台型公司 (Google) 与终端品牌 (OEM) 之间的微妙权力博弈——架构层面的细微设计选择, 决定了价值链中谁拥有”屏幕主权”。

4.3.3 Waymo 的”可证明安全” AI 架构 Waymo 对架构设计的终极信仰是: 在自动驾驶领域, 安全绝不能作为功能的”事后补救层”, 而必须从架构的胚胎期便深入基因骨髓。

Waymo Foundation Model 生态: 基座是宏大的 Waymo 基础模型, 其上承载三个核心功能组件——

- **Driver (驾驶员模型):** 实际执行驾驶决策;
- **Simulator (仿真器引擎):** 生成训练与测试场景;
- **Critic (评估批评模型):** 严苛地评分与挑战 Driver 的决策。

三大组件在相同的底层 AI 架构上并行生长，形成不断自我强化的闭环。

System 1 / System 2 双轨混合架构：

- **传感器融合编码器 (System 1, 快思考)**：以极低延迟融合激光雷达、摄像头、雷达的海量实时数据，通过神经嵌入直接输出周边物体的运动属性与语义信息，满足绝大多数情况下的本能避撞与快速反应；
- **驾驶视觉语言模型 (Driving VLM, System 2, 慢思考)**：与 Gemini 大语言模型深度融合，专门负责慢思考。当车辆遭遇极其罕见、甚至超出人类常识的长尾场景（前方道路上出现燃烧的油罐车、穿戴奇异服饰的游行队伍）时，**System 2 被激活**，利用语言模型内蕴的广泛物理常识进行深度逻辑推理，从而指挥车辆执行超越单纯几何避障规则的高级语义避险行为（立刻掉头、规划数十公里外的绕行路线）。

Teacher-Student 知识蒸馏机制：为了在资源受限的车端计算平台上实现这一庞然大物的实时运行——

- **Teacher Driver / Teacher Critic** 在云端算力中心进行慢但深思熟虑的博弈与仿真推演；
- 优质应对策略与隐式世界认知知识被**压缩与蒸馏**到极度轻量化的 **Student Driver** 小模型；
- Student 在实体车辆芯片上完成低延迟实时驾驶推理；
- 执行前通过专门的**独立车载验证层**对生成轨迹的合法性进行最终物理拦截。

双循环强化学习：

- **Inner Loop (内部强化)**：Simulator 与 Critic 在虚拟空间内对 Driver 进行百万倍速的残酷试炼；
- **Outer Loop (现实循环)**：真实路测数据捕捉 Driver 的次优表现，在云端寻找改进对策并再次下发更新。

结构化世界表征：与纯黑盒端到端网络不同，Waymo 基础模型同步生成结构化的世界表征（交通参与者对象、明确的语义属性、道路拓扑图），这为推理时刻的**安全验证准则**提供了关键的可视化与可解释性抓手。

Waymo 第 6 代 Driver（2024 年 8 月发布，2026 年 2 月 12 日全自主上路）：传感器数量较第 5 代减少 42%、BOM <2 万美元（成本降 >50%）；雪天验证完成（波士顿、匹兹堡、丹佛）；跨平台部署——Zeekr 联合开发的 Ojai 专用车 + 现代 IONIQ 5（**5 万辆订单，史上最大 AV 单笔订单**）。

运营指标（2026 年初）：2025 年付费打车 1,500 万次（4× 2024）；累计 2,000 万 +；周打车约 40 万次；车队约 2,500 辆；累计 2 亿英里全自动驾驶；运营 Phoenix、SF、LA、Austin、Miami；扩展华盛顿 DC、圣地亚哥、东京（测试）、伦敦（2026 公共）、纽约（测试）。

4.3.4 Waymo vs Tesla: 两种自动驾驶架构哲学的根本对立

维度	Waymo	Tesla
传感器	LiDAR + 相机 + 雷达 + 音频 (多冗余)	纯视觉
高精地图	城市级别必需	不依赖
ODD	地理围栏 L4	无围栏 L2++ → Cybercab L4
责任	Waymo (车辆所有者)	驾驶员 (消费车) / Tesla (Cybercab)
BOM	第 6 代 <2 万美元	1—2 千美元 (仅传感器)
软件	传感器融合 + 世界模型 + System1/2	端到端 NN
扩张哲学	城市逐个 HD 建图 “今天就比人安全”——硬件冗余优先	OTA 全球一夜部署 “数据 + 规模会收敛”——软件优先

会收敛吗? 不太可能。Waymo 第 6 代将传感器减 42% 显示 ML 成熟度允许减冗余; Tesla 验证舰队在凤凰城被发现搭载 LiDAR (用于地面真值标注, 非消费部署)。但 Waymo 的地理围栏 + L4 责任模型从根本上需要冗余传感器认证“无司机”声明; Tesla 的经济模型 (消费车 + Optimus) 无法承担 LiDAR + HD 地图, 必须通过端到端 ML 规模化弥合差距。**两种范式各自收敛的临界点最可能落在 2027—2029 年——届时端到端 VLA 在仅相机配置下能否提供可证明的安全性, 将决定未来产业版图。**

1.6.4 4.4 平台赋能型的其他亚类

平台赋能型内部进一步细分为四个亚类, 每一类对应不同的“关键瓶颈”控制策略:

亚类	代表	核心赋能资产
开放 OS 平台型	Google (AAOS)	Android/AAOS/Wear OS 生态、Treble HAL、Mainline 模块化更新
自动驾驶平台型	Mobileye	EyeQ 异构 SoC、REM 众包地图、RSS 形式化安全模型、True Redundancy

亚类	代表	核心赋能资产
算力 + 仿真平台型	NVIDIA	DRIVE + Hyperion + Halos; Isaac + Cosmos + GR00T; 训练-仿真-部署”三计算机”范式
量产工程平台型	Bosch、Aptiv	中央车载计算 + 区域控制 (Aptiv SVA 的 PDC + Open Server Platform)、全栈量产工程

4.4.1 Mobileye: 以形式化安全为核心的自动驾驶平台 Mobileye 代表了一种独特的架构哲学：不追求拥有终端品牌，而是通过深度技术专精成为自动驾驶领域的”安全方法学事实标准”。其核心赋能资产是三位一体的——

- **EyeQ 异构 SoC 路线图**: EyeQ4 (2.5 TOPS, 2018) → EyeQ5 (24 TOPS, 2021) → **EyeQ6 Lite/High** (2024 年末量产, 单芯支持完整 ADAS 至 L2++) → **EyeQ Ultra** (12 个双线程 RISC-V 核 + 256 GFLOPS Arm GPU + 64 个专有加速核, <100W TDP, 目标单芯片 L4, 2025—26 量产)。EyeQ 系列的差异化不在峰值算力, 而在**每瓦特性的极致优化**——这一指标对量产嵌入式系统至关重要。
- **REM (Road Experience Management) 众包高精地图**: 通过数百万辆量产车的视觉数据实时回传, 构建并持续更新厘米级 HD 地图。截至 2025 年, REM 覆盖超过 250 亿公里数据, 是全球规模最大的众包驾驶数据集之一。
- **RSS (Responsibility-Sensitive Safety) 形式化安全模型**: Mobileye 在 2017 年发表的开创性论文中提出 RSS——一套基于数学公理 (安全纵向距离、安全横向距离、合理假设、谨慎避撞、责任分配) 的形式化驾驶决策框架。RSS 不是黑盒 NN, 而是**可证明的安全包络**。在端到端神经网络日益成为产业主流的今天, RSS 作为”运行时安全监视器”被越来越多地与黑盒模型组合使用——这与 Waymo 的”独立车载验证层”思路异曲同工。

Mobileye 还将这一平台能力扩展到非汽车领域。2025 年其宣布将 EyeQ 系列芯片与感知/规划栈延展至人形机器人——这一举措与 Tesla “FSD 栈赋能 Optimus”的逻辑高度类似, 但 Mobileye 走的是**赋能其他机器人本体厂商**的水平路径而非垂直整合。

架构哲学的内在张力: Mobileye 的优势在于自动驾驶域的极致专精与可证明安全, 但它的劣势也根植于此——**不像 Android/Tesla 拥有大消费端 OS 生态**。这使得 Mobileye 的价值天花板受限于”被集成”的角色——它无法直接拥有终端用户关系, 必须通过 OEM 客户间接触达消费者。

4.4.2 NVIDIA：训练-仿真-部署的”三计算机”范式 NVIDIA 是平台赋能型中最具雄心的玩家。其核心架构判断是：**从云到车/机器人的统一算力与开发栈是这场革命的最大赋能机会。**这一判断转化为”三计算机”（Three Computers）范式：

- **训练计算机**：数据中心 GPU（H100、B200、GB200 NVL72）+ DGX/HGX 集群 + CUDA + cuDNN + Megatron + NeMo 训练栈；
- **仿真计算机**：Omniverse + Isaac Sim + **Cosmos 世界模型**（已被 Toyota DriveOS、1X、Skild AI、Figure AI、Agility、Uber、Waabi、Foretellix、小鹏、Galbot、Fourier 同时采用）；
- **运行时计算机**：DRIVE Thor（车载，2,000 TFLOPS FP8）+ Jetson Thor（机器人，2,070 TFLOPS FP4）+ DRIVE OS + Isaac ROS。

横跨三个计算阶段的统一开发栈是 NVIDIA 的核心壁垒——开发者可以在 Omniverse 中训练并验证一个 VLA 模型，然后无缝部署到 Drive Thor 或 Jetson Thor，**所需的代码与工具链改动极小。**这一”一次开发、跨形态部署”的能力是当前其他平台难以匹敌的。

NVIDIA 进一步通过 **Halos（自动驾驶安全平台）**与 **GR00T（机器人基础模型）**把方法论框架与基础模型一并纳入赋能矩阵。其商业模式不卖整车、不卖整机，仅卖平台、芯片、工具链、模型——但通过**赋能整条产业链**实现价值放大。

架构哲学的内在张力：NVIDIA 平台极强，但车企/机器人公司仍需自行完成产品定义与责任闭环。NVIDIA 不为终端产品的安全声明、法规合规、用户体验直接负责——这既是它的优势（轻资产、高利润），也是它的天花板（始终是”被集成”而非”集成方”）。

4.4.3 Bosch、Aptiv：量产工程的平台型 Tier-1 Bosch 与 Aptiv 代表了平台赋能型中最”传统”也最具量产工程深度的一支。他们从 Tier-1 起家，但在 SDV 时代主动重塑自身定位——从”卖单个 ECU”转向”卖中央车载计算 + 区域控制 + 全栈工程服务”。

Bosch 的转型路径：从经典 Tier-1 向”软件定义移动出行公司”演化。其差异化资产是车身、底盘、传感器、动力总成、软件、AI 的全栈集成能力。在中央车载计算与区域控制方向，Bosch 押注于**跨域控制（cross-domain control）**范式——一个 HPC + 多个 ZCU 配合 Bosch 的全栈量产工程能力，覆盖从 ADAS 到底盘控制的关键安全功能。AUTOSAR 2026 大会的主题转向”开源、社区源与商业汽车软件的结合”——这一信号背后是 Bosch 等头部 Tier-1 对开放生态的战略接纳。

Aptiv 的 Smart Vehicle Architecture (SVA)：Aptiv 提出的 SVA 是当前最具代表性的 Tier-1 区域化架构方案。其核心是 **PDC（Power Data Center）**区域控制器 + **OSP（Open Server Platform）**开放服务器平台的双层组合。PDC 负责本地 I/O、供配电、低速通信；OSP 作为高算力中央计算节点，硬件可交换、软件抽象化，使得不同车型可以根据需要灵活配置算力梯度。Aptiv 的核心战略意图是**延长车辆生命周期并平滑算力升级**——即通过架构层面的解耦，使一辆车在 10—15 年生命周期内可以更换中央计算单元而不重做整车线束。

架构哲学的内在张力：Bosch、Aptiv 这类量产工程平台型公司平台能力极强，但**品牌触点与数据闭环通常掌握在 OEM 手中**。最终产品定义仍需 OEM 与上层软件伙伴共同完成。这种”价值在中间、控制权在两端”的位置是传统 Tier-1 在 SDV 时代最深层的战略焦虑。

4.4.4 四类平台赋能型的共性与差异 这四类平台赋能型公司的共性是不追求拥有终端品牌，而是通过赋能终端品牌实现价值最大化。差异在于它们掌握的”关键瓶颈”不同：

- **Google**（开放 OS 平台型）掌握的是**消费操作系统与开发者生态**——这是消费侧的入口控制权；
- **Mobileye**（自动驾驶平台型）掌握的是**自动驾驶算法与可证明安全方法学**——这是责任侧的认证控制权；
- **NVIDIA**（算力 + 仿真平台型）掌握的是**统一算力底座与跨形态开发栈**——这是开发者侧的生产力控制权；
- **Bosch、Aptiv**（量产工程平台型）掌握的是**大规模量产工程与跨域系统集成**——这是供应链侧的工程控制权。

这一谱系揭示了一个深层规律：**在复杂系统产业中，价值链上的不同”控制点”可以同时被不同主体掌握而不必合并**。我们将在第六章系统化这一”架构控制点”概念。

1.6.5 4.5 三类原型的横向对比矩阵

图 4.6 代表性 OEM 五维架构能力雷达图 (Snapshot 2026.1.31, 各维满分 5)。

第四章前三节分别深入剖析了三类架构原型——垂直闭环型 (Tesla)、跨设备生态型 (华为与小米)、平台赋能型 (Google 与 Waymo) ——的哲学根基、复用机理与限制条件。本节将这三类原型置于同一张矩阵中, 进行系统化的横向对比。这一对比的目的不在于评判孰优孰劣, 而在于揭示每一类原型背后的资源禀赋假设、组织能力路径依赖、监管环境适配, 以及由此衍生的可行性边界。

4.5.1 对比矩阵的设计原则 跨原型对比需要避免两种常见陷阱。第一种陷阱是**简单归纳**, 即将所有原型化约为一个单一维度 (如”开放 vs 封闭”或”软件优先 vs 硬件优先”), 这种简化遮蔽了实际架构决策的多维度耦合。第二种陷阱是**价值预设**, 即默认某一类原型为”先进”或”落后”, 这种预设无视各原型在不同市场、不同监管、不同资源条件下的具体合理性。

本节采用的对比矩阵基于五个独立维度——架构控制深度、复用机理、生态边界、监管适配、扩展瓶颈——每个维度对每类原型给出客观描述, 不做高低排序。

4.5.2 五维度对比矩阵 表 4.5: 三类架构原型横向对比矩阵

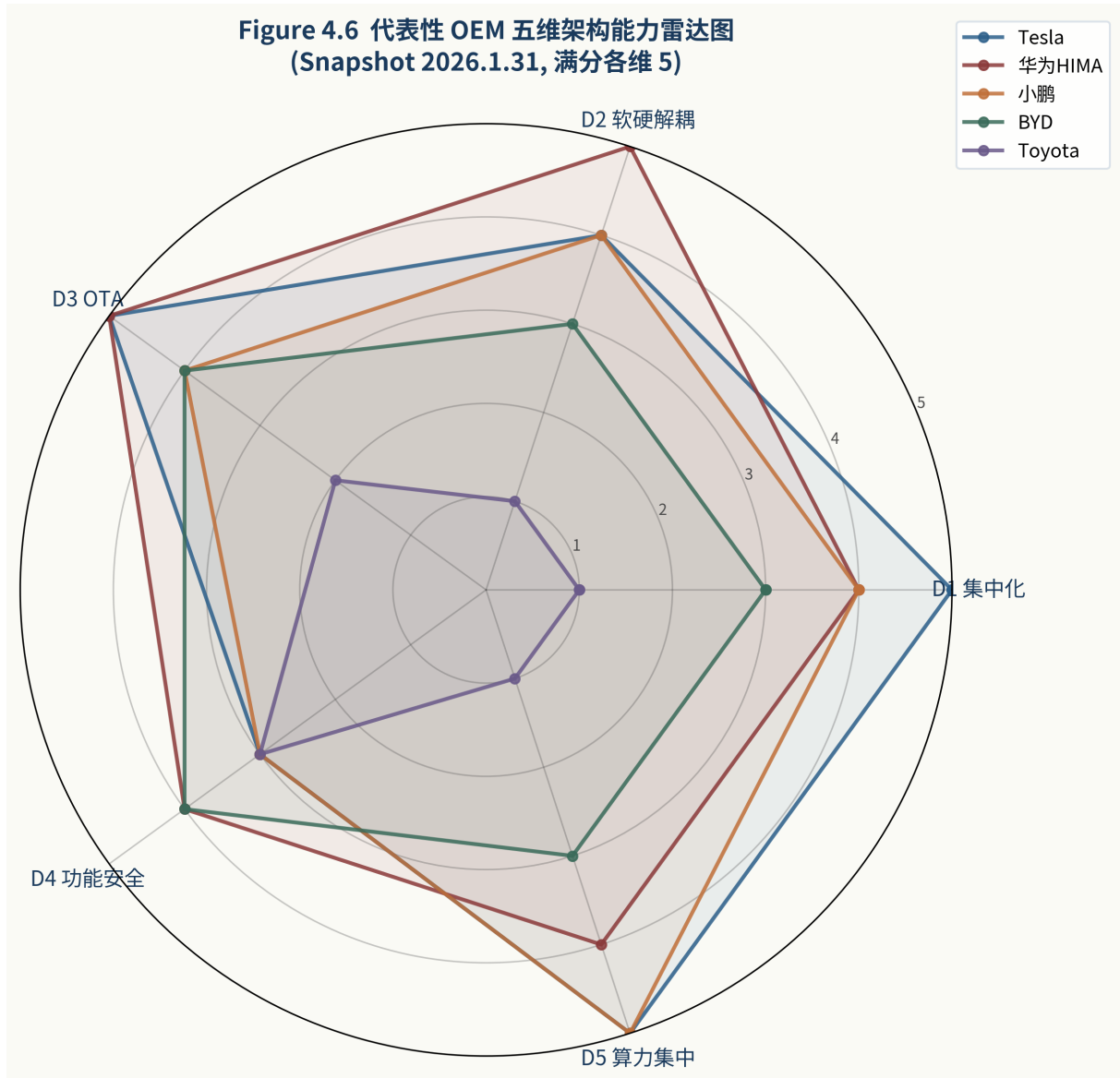


Figure 5: 图 4.6 代表性 OEM 五维架构能力雷达图 (Snapshot 2026.1.31, 各维满分 5)。

维度	垂直闭环型 (Tesla)	跨设备生态型 (华为 / 小米)	平台赋能型 (Google / Waymo)
架构控制深度	全栈自有——硅片、OS、应用、云、机器人全部自研	关键层自有 + 外围合作——内核 OS、基础模型、AI 芯片自研, 整车制造可选自有 (小米) 或合作 (华为 HIMA)	关键 IP 自有 + 终端合作——TPU、Gemini、Android、AAOS 自有, 终端形态依赖合作伙伴
复用机理	同源芯片 + 同源感知栈 + 同源训练基础设施 + 团队合并	同源微内核 + 同源大模型 + 同源算力底座; 硬件可异构 (小米路线)	同源芯片底座 + 同源 LLM 基础设施; 终端架构高度异构
生态边界	封闭——无第三方应用生态, 应用创新受限于单一研发节奏	半开放——HarmonyOS 开发者生态 1,000 万 +; HyperOS 围绕小米硬件生态	高度开放——Android 全球生态 + TPU 外部客户 (Anthropic、Apptronik 等)
监管适配	风险高——纯视觉路线在 L4 责任声明上仍受监管质疑; 端到端 NN 的可解释性挑战	中等——中国监管环境适配优秀; 海外市场面临数据主权挑战	优秀——Waymo 拥有最完整的功能安全工程方法学; Google 在 GDPR/PIPL 合规上久经考验
扩展瓶颈	资本与人才——同时垂直整合所有层需要数百亿美元投入; 任何单层失败将连锁影响全栈	政治与地缘——华为受芯片制裁倒逼; 小米受全球供应链与市场准入约束	终端控制点缺失——Google 不直接拥有汽车与机器人终端, 依赖合作伙伴的执行能力
典型资本量级	数百亿美元级 (仅 Tesla 一家市值数万亿美元支撑这一战略)	百亿到数百亿人民币级 (华为研发投入与小米生态投入)	数百亿美元级, 但分摊到多产品线
典型时间窗	长期主义——FSD 项目自 2014 年至今 12 年仍在演化	中长期——HarmonyOS 自 2019 年至今 7 年; 小米造车从 2021 年宣布到 SU7 量产仅 3 年	长期——Waymo 自 2009 年 Google 自动驾驶项目至今 17 年

维度	垂直闭环型 (Tesla)	跨设备生态型 (华为 / 小米)	平台赋能型 (Google / Waymo)
失败时的损失模式	全栈失败——任何一层崩溃可能拖垮整个体系；难以局部止损	模块化失败——某一形态失败不致命, 可专注其他形态	控制点失败——若 TPU 或 Gemini 失去竞争力, 平台地位崩塌；终端层的失败不直接致命
代表性成功指标	FSD 量产规模 + Optimus 跨形态复用证明	HarmonyOS 装机量 + 华为 HIMA 装车数 + 小米 SU7 销量	TPU 外部客户数 + Android 全球市占 + Waymo 商用 Robotaxi 里程
典型限制案例	纯视觉路线的安全争议、Cybertruck 等” 虚荣项目” 消耗、CEO 个人风险	华为受芯片代际封锁、小米首款车型质量与品牌挑战	Apple 终端缺失 (Google 进入消费硬件的反复失败)、合作伙伴执行力差异 (Pixel 与三星生态分歧)

4.5.3 不可比拟性与可借鉴性 上述对比表的一个关键洞察是：这三类原型彼此之间不可简单替换，但其内部的具体实践却可以选择性借鉴。

不可替代性的来源在于**资源禀赋与组织 DNA 的根本差异**。一家传统 OEM 不可能在 18 个月内”变成 Tesla”，因为垂直闭环要求的资本规模、人才密度、组织扁平度都是 20 年累积的产物；一家美国大学衍生公司不可能”变成华为”，因为水平微内核 + 大模型平台要求的国家级研发协调能力和长期忍耐力是另一种文明传统的产物；一家中国互联网公司不可能”变成 Google”，因为联邦 AI 平台要求的全球开源治理经验、跨监管区域合规能力、与硬件厂商的长期信任关系都不是短期可建立的。

可借鉴性的来源在于**具体实践与决策模式的局部可复制性**。其他玩家可以借鉴 Tesla 的”形态无关感知栈”思路，将自家的视觉算法或控制算法跨形态复用；可以借鉴华为的”水平微内核 + 大模型平台”思路，在某一个细分领域（如车-手机协同）做小规模实现；可以借鉴 Google 的”开放平台 + 顶级 IP”思路，将自家的核心技术开放给生态合作伙伴，形成局部的”小生态”。

具体而言，本研究观察到以下三类”中间路径”实践：

路径一：垂直整合的局部化。一些 OEM 选择在某些关键层（如 OS、电池管理、ADAS）做垂直整合，在其他层（如座舱、信息娱乐）继续采购或合作。Rivian 在中央计算上的极致简化、比亚迪在电池管理上的垂直整合、蔚来在自研芯片（神玑 NX9031）上的投入都是这一路径的体现。

路径二：跨设备生态的窄域化。一些非超级生态玩家选择只在特定的设备-设备协同上做生态

化,而非全栈跨设备。例如 Polestar 与 Volvo 在车-手机协同上的合作、Mercedes 在 MB.OS 内部统一座舱与 ADAS 的尝试。

路径三：平台赋能的逆向化。一些 OEM 选择”反向赋能”——不做平台主,但也不做平台消费者,而是将自家的核心能力(如 BMW 的高级 ADAS、丰田的混动控制系统)授权给其他玩家,形成局部的”小生态”。

4.5.4 选择原型的决策框架 对于产业界读者,本节给出一个简化的决策框架,以辅助判断哪一类原型最适合参照。这一框架包括四个判断条件:

条件一：资本与人才可达性。垂直闭环需要 100 亿美元以上的累积投入与 5,000 人以上的顶级工程师团队;跨设备生态需要至少一个超级终端入口(如手机或智能家居)与 100 亿美元级的研发投入;平台赋能需要全球级的开源治理能力与 200 亿美元以上的技术 IP 储备。

条件二：监管适配可控性。垂直闭环型在欧美高度成熟监管市场风险较高,在中美等大单一市场风险中等;跨设备生态型在中国市场风险较低,在欧美面临数据主权挑战;平台赋能型适应几乎所有监管环境,但代价是单点利润率较低。

条件三：组织 DNA 兼容性。垂直闭环要求扁平、快速、容错的硅谷创业文化;跨设备生态要求超长期主义、强中央协调、忍耐技术封锁的能力;平台赋能要求成熟的开放治理、跨文化协调、与硬件合作伙伴的长期信任。

条件四：失败可承受性。垂直闭环的失败可能致命,因为全栈耦合;跨设备生态的失败可局部止损;平台赋能的失败局限于单一控制点。这一维度对资源有限的玩家尤其重要——选择失败成本可承受的原型,比追求最大估值天花板更重要。

需要强调的是,这一决策框架不能替代具体的战略判断。每家公司的实际选择都受其历史路径、现有资产、领导团队偏好、市场窗口的综合影响。本研究的目的是提供分析工具,而非给出具体决策。

4.5.5 三类原型的协同与竞争 最后值得指出的是:三类原型之间不仅是竞争关系,也存在显著的协同空间。

Tesla 的开放节点包括接受 NVIDIA Cosmos 世界模型作为训练补充、与某些云厂商合作做训练数据备份、向部分研究机构开放 FSD 数据集(虽然规模有限)。这意味着即便最封闭的垂直闭环型,也无法做到完全的零外部依赖。

华为的开放节点包括 OpenHarmony 开源(不同于 HarmonyOS NEXT 的全自研版本)、HIMA 联盟中对鸿蒙智行品牌的多 OEM 共用、Pangu 大模型在政府与企业客户中的对外开放。这意味着即便垂直整合最深的跨设备生态型,也保留了相当的对外赋能空间。

Google 的封闭节点包括 TPU 的内部优先使用(Gemini 3 完全在 TPU 上训练)、Pixel 与 AAOS 的内部协同、Waymo 商业模型对外部数据合作的限制。这意味着即便最开放的平台赋能型,也存在不可分享的核心控制点。

这种”既封闭又开放”的混合模式,反映了当代复杂系统架构的真实形态——纯粹的封闭与纯粹的开放都不可持续,胜出的玩家都在持续优化”开放什么、封闭什么、开放给谁、何时开放”

的精细判断。本研究第六章将进一步系统化这一观察, 引入”七大架构控制点”作为分析工具。

1.7 第五章架构成熟度的统一框架：AR0—AR5 与 AI²-ML

1.7.1 5.1 启发与方法论

SAE J3016 自动驾驶分级 (L0—L5) 的成功在于它提供了一个跨企业、跨技术路线、跨监管区域的共同语言。任何关于自动驾驶能力的讨论，都可以快速锚定在 L2/L3/L4/L5 上。这种“分级语言”对架构演进同样必要——但当前业界缺乏类似的统一框架。

机器人行业曾尝试类似分级 (EarthSense 的 Hands off / Eyes off / Mind off / Monitoring off / Development off) 但局限于野外/农业机器人，未形成全产业共识。汽车 E/E 架构方向上，博世六阶段框架影响最大，但其本质是单一产业内的演进路径，并未抽象到可跨域使用的层面。

本研究提出 **AR0—AR5 架构能力门槛框架** (AR = Architecture Readiness)，尝试在更高抽象层次上统一汽车 E/E、机器人、广义智能终端三大领域的架构演进。**重要的方法论原则是：AR 级别表征能力门槛，而非时间线；低阶不会随高阶出现而消失，而会与高阶长期共存。**这与 SAE J3016 的精神一致——道路上 L0/L1/L2/L3 车辆将长期共存数十年。

1.7.2 5.2 AR0—AR5 架构能力门槛框架

AR0 —机电孤岛 (Mechanical-Electrical Island) 核心特征：功能一对一控制器，软硬件深度耦合，软件随硬件走，几乎无远程更新能力。

关键指标：ECU/控制器数量高、跨域协同弱、OTA 范围窄或不存在。

典型状态：传统分布式燃油车、单体工业机器人控制器、早期智能手机的离散基带/AP/ISP/NPU 时代。

架构含义：此阶段属于“数字石器时代”。系统对外部物理世界缺乏全局感知能力，本质上仍是执行器集合，不具备现代意义上的可演化平台属性。

AR1 —域级集成 (Domain-Level Integration) 核心特征：座舱/车身/智驾等域整合，AUTOSAR Classic/Adaptive 并存或域内部分采用 Adaptive，OTA 开始规模化但仍以单 ECU 为单位。

关键指标：5±2 个域控制器，跨域通信仍受限于物理总线带宽。

典型状态：2024—2026 多数量产 SDV 过渡态、协作机械臂、需要遥控介入的人形机器人演示。

架构含义：出现了 SOA 雏形与基础中间件。但域与域之间的信息流转仍被严格切断，整车或系统无法产生涌现式的联合智能决策。

AR2 —区域化平台 (Zonal Platform) 核心特征：区域控制 + 中央计算，服务化诊断，虚拟化与标准信号目录。

关键指标：控制器数量明显下降 (典型 7—15 个)，整车软件平台化，跨车型/跨产品线复用增强。

典型状态：高端 EV、新一代 SDV 平台、Boston Dynamics Spot 商用产品、Agility Digit。

架构含义：Hypervisor 虚拟化让消费级 OS 与硬实时微内核在同一异构 SoC 上安全共存——这是 AR 演进中最具转折意义的技术突破。

AR3 —跨设备协同智能体 (Cross-Device Collaborative Agent) 核心特征：车/手机/穿戴/家居/机器人共享身份、上下文与任务迁移。

关键指标：跨设备任务连续性、统一账号与策略、边云协同调度、跨形态算力池化。

典型状态：华为/小米生态方向，Google/Android 也在向车内外联动扩展。

架构含义：架构哲学从“单一设备的最优化”转向“用户身份与任务流的最优化”。设备本身变成可替换的算力 + 传感器组合，软件生态从设备级跃迁到生态级。

AR4 —多体物理 AI 平台 (Multi-Embodiment Physical AI Platform) 核心特征：训练、仿真、部署闭环统一；世界模型与基础模型跨车/机器人共享。

关键指标：仿真与真实数据闭环效率、模型复用率、策略迁移效率、跨形态泛化能力。

典型状态：Waymo Foundation Model、NVIDIA Cosmos + GR00T、Tesla AI5+FSD V14+Optimus 同源栈、Mobileye Robotics 扩展。

架构含义：架构彻底打破底层硬件束缚。分布式微内核 OS + 端到端基础模型实现跨智能手机、穿戴、车机、机器人之间的算力动态池化共享。系统普遍引入 System 1（瞬时直觉感知）与 System 2（大模型逻辑推理）双轨机制。

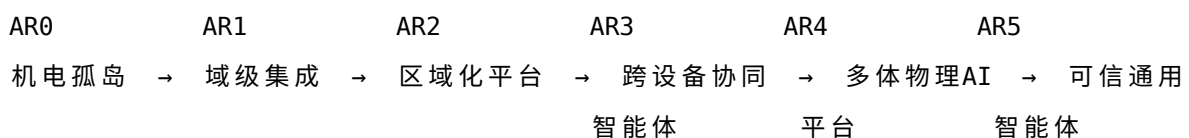
AR5 —可信通用智能体 (Trusted General Embodied Agent) 核心特征：通用策略层可跨多种实体执行，持续证明安全/合规/责任。

关键指标：不只是“能做”，而是“可证明地做”。

典型状态：目前尚无成熟量产范式，处于研究与原型阶段。

架构含义：系统步入“硅基生命体”的门槛。仅凭一套底层多核分发微内核 OS + 一个洞悉世界物理定律的基础模型，便能同时驱动用户的随身全息设备、暴风雪山区的无人物理穿梭舱、厨房颠勺炒菜的双足人形机器人。计算环境彻底去中心化与泛在化。跨设备、跨行业的物理隔离壁垒全线崩塌。

1.7.3 5.3 AR 分级的可视化



需要再次强调：**这是能力门槛而非时间线**。在 2030 年代，AR0—AR2 仍将作为传统车型与工业控制系统的主流形态长期存在；AR3—AR4 将成为高端智能终端的竞争前沿；AR5 仍

处于研究与原型阶段。

1.7.4 5.4 引入 Architecture Intelligence Maturity Levels (AI²-ML)

AR 框架刻画的是**被设计系统**（车辆、机器人、智能终端）的架构成熟度。但架构本身作为一项工程活动，也有其方法论与工具支持的成熟度。我们将这一维度称为 **Architecture Intelligence Maturity Levels (AI²-ML)**，即”架构智能成熟度等级”——衡量一个组织在做架构决策时所能调用的方法论、工具、知识库、AI 辅助的成熟度。

AI²-ML 的 5 级定义

级别	名称	核心特征	典型工具
L0	Drawing (绘图)	用通用绘图工具表达架构，无形式化语义	Visio、Miro、PowerPoint
L1	Modeling (建模)	用建模工具表达架构，有形式化语义但缺乏约束检查	PREEvision、Capella、Cameo、Enterprise Architect
L2	Constraint-Checked (约束检查)	模型支持约束规则与一致性检查	PREEvision Pro、Capital Systems Architect
L3	Reference-Aware (参考感知)	工具内置参考架构库，能基于对标进行架构评估	参考架构数据库、行业对标平台
L4	AI-Assisted (AI 辅助)	工具用 AI 辅助架构决策、自动布局、约束推理	AI 增强的架构平台
L5	Autonomous (自主)	工具能在给定目标与约束下自动生成与演化架构	尚未出现成熟产品

AI²-ML 的 5 维评分维度 AI²-ML 框架在评估工具或方法论时，使用以下 5 个独立维度：

- **D1: Architectural Centralization (架构集中化程度)** ——支持从分布式到中央计算的全谱表达；
- **D2: Software-Hardware Decoupling (软硬解耦深度)** ——支持服务化、虚拟化、抽象层的建模；
- **D3: OTA Maturity (OTA 成熟度)** ——支持全生命周期更新管理的建模；

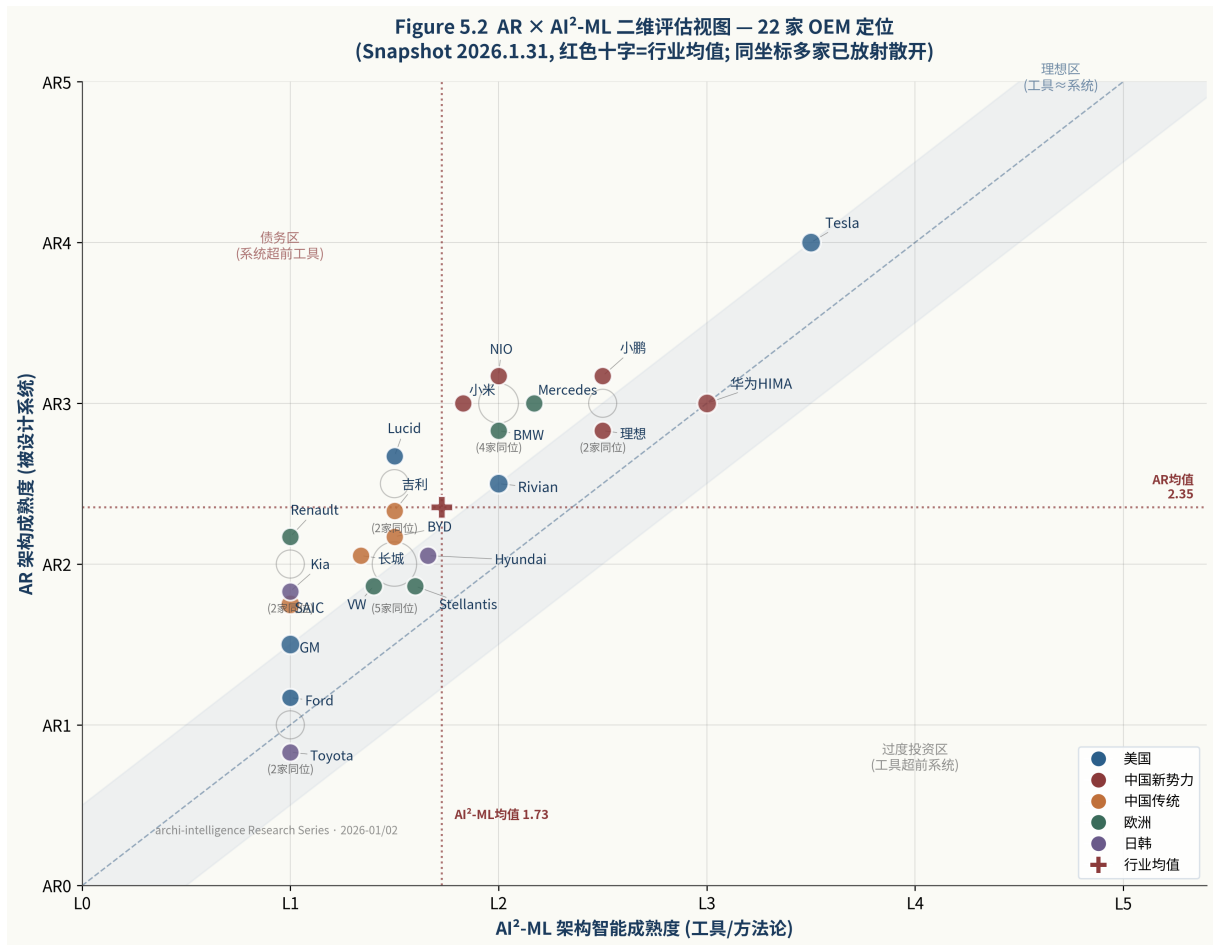


Figure 6: 图 5.2 AR × AI²-ML 二维评估视图——22 家 OEM 定位 (Snapshot 2026.1.31)。对角线带为理想区 (工具 ≈ 系统)。

- **D4: Functional Safety Architecture (功能安全架构)** ——支持 ISO 26262/SO-TIF/ISO/PAS 8800 的形式化表达;
- **D5: Compute Concentration (算力集中度)** ——支持异构计算资源的拓扑与调度建模。

AR × AI²-ML 二维评估视图 图 5.2 AR × AI²-ML 二维评估视图——22 家 OEM 定位 (Snapshot 2026.1.31)。对角线带为理想区 (工具 ≈ 系统)。

AR 与 AI²-ML 是两个正交的独立维度。一个组织可能拥有 AR4 级别的系统 (Tesla FSD + Optimus)，同时拥有 L3—L4 级别的内部架构工具；也可能拥有 AR2 级别的系统，但因为依赖 L1 级别的绘图工具，使得架构演化举步维艰。

这一二维视图揭示了架构决策质量的真实约束：单纯追求系统架构成熟度 (AR) 而忽视架构智能工具 (AI²-ML) 的同步进化，会导致“架构债务”的指数级累积。一个 AR4 系统如果用 L1 工具维护，其演化成本将超出团队承受范围。这正是 CARIAD 困局的深层根源之一——目标架构是 AR3—AR4 级别的 SSP/E3 2.0，但工具与方法论仍停留在 L1—L2 水平。

AI²-ML 框架的跨产业适用性 AI²-ML 框架并非主张普适性。不同产业由于其约束特征、生命周期、监管强度的差异，对系统级架构工具的需求强度也截然不同。本研究对六大产业的适用性做了初步评估：

产业	适用性评分	关键判断依据
汽车 E/E 架构	5 / 5	长生命周期（10—15 年）+ 严苛监管（ISO 26262、UN R155/R156、ISO/PAS 8800）+ 多供应商协作 + 跨车型复用 + 软硬解耦演进——所有 5 个维度都强约束驱动
机器人 / 具身 AI	3 / 5	当前产业仍处早期阶段，标准与监管框架不完整，工具链分散。预计 24—36 个月后随产业成熟度上升而需求显化
航天系统	5 / 5	极长生命周期（15—30 年）+ 极严监管（DO-178C、ECSS、NASA NPR）+ 极低容错预算。市场规模小但适用性极高
工业自动化	4 / 5	中长生命周期 + IEC 61508/62443 监管 + 模块化设计需求。略低于汽车的原因是单体系统复杂度通常较低
智能手机 / 消费电子	2 / 5	迭代速度快（年度发布节奏）+ 单一品牌闭环 + 标准化 SoC 平台主导——架构决策更依赖 EDA 工具链而非系统级架构工具
医疗器械	4 / 5	监管严苛（FDA、CE MDR、IEC 62304）+ 长验证周期 + 可追溯性要求高。略低于汽车的原因是软硬件解耦演进的产业紧迫性不及汽车

图 5.3 “跨行业”工具-系统”匹配模式对比。汽车因系统受监管/安全驱动持续演进、但专用架构工具稀缺，点偏债务区、过度投资区为空；航天因极强监管强制工具与系统同步而贴近

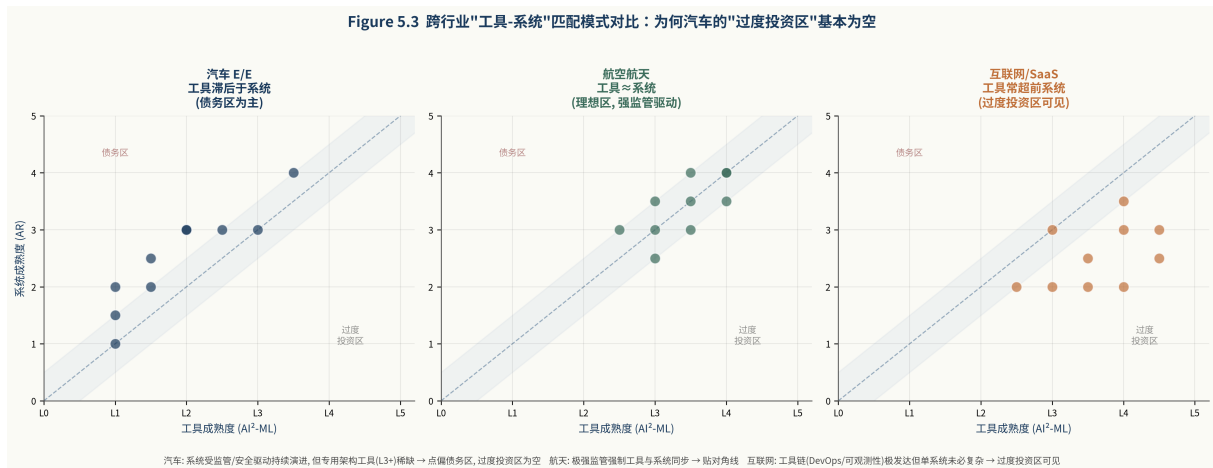


Figure 7: 图 5.3 跨行业”工具-系统”匹配模式对比：为何汽车的”过度投资区”基本为空

理想区；互联网因工具链发达但单系统未必复杂，过度投资区可见。

这一评估的核心启示是：**AI²-ML 框架最具价值的应用领域是那些同时满足三个条件的产业——长生命周期、强监管约束、多供应商协作下的软硬件解耦演进。**汽车 E/E 架构是这一交集的”完美样本”，因此本文以汽车为锚点展开方法论讨论。在适用性较低的行业（如智能手机），系统级架构工具的边际价值递减——这并非框架的失败，而是产业约束特征的客观差异。

AI²-ML 与现有方法论的关系 AI²-ML 框架不寻求替代现有的架构方法论，而是希望在更高抽象层次上提供一个关于”架构工具与方法论本身的成熟度”的评估视角。它与以下框架的关系是互补的：

- **MBSE / SysML:** MBSE (Model-Based Systems Engineering) 与 SysML 是 L1—L2 级别的核心方法论支撑。AI²-ML 不取代它们，而是评估它们在具体组织内的实施深度；
- **TOGAF / Zachman:** 企业架构框架主要服务于业务-IT 对齐。AI²-ML 聚焦于工程系统架构，与企业架构框架属于不同层面；
- **ASPICE / CMMI:** 过程能力成熟度模型关注的是开发流程的成熟度。AI²-ML 关注的是工具与方法论的成熟度——前者问”流程做得多规范”，后者问”工具支持多智能”。

这一定位使 AI²-ML 框架既不与现有标准冲突，又能填补”架构工具成熟度评估”这一长期空白的方法论生态位。

1.7.5 5.5 AR 与 SAE J3016 的关系

需要明确：AR 框架与 SAE J3016 不是相互替代，而是相互独立的两个维度。

- **SAE J3016** 描述的是车辆自动化能力——在特定 ODD 内，车辆能否替代人类驾驶员；
- **AR 框架**描述的是承载这种能力的架构成熟度——系统在多大程度上是软件定义的、可演化的、可跨形态复用的。

理论上, 一辆 L3 自动驾驶车可能由 AR1 架构实现 (高度定制的孤立 ADAS 系统), 也可能由 AR4 架构实现 (端到端基础模型驱动)。在产业实践中, **L4/L5 自动驾驶事实上要求 AR3 以上的架构基础**——因为持续舰队学习、多场景泛化、跨车型复用都依赖跨设备/跨形态的架构能力。

1.8 第六章总结、研究边界与开放问题

1.8.1 6.1 核心论点综述

本文的核心论点可以浓缩为以下六条：

论点一：架构是关于权衡的工程艺术，而非关于完美的科学。 两条底层定律——Trade-off 与 Why > How——跨越两千年依然有效。当代复杂系统架构决策的不可逆性，使得 Why 的论证质量直接决定系统的长生命周期成本。

论点二：当代各产业架构正经历”基础设施趋同、控制语义与证明义务分化”的双层演化。 共用底座（异构 SoC、虚拟化、SOA、世界模型、OTA）正在跨产业收敛；行业特定上层（失效哲学、安全壳、责任归属、合规框架）将继续分化并因监管强化而进一步固化。

论点三：失效哲学（fail-soft vs fail-operational）是汽车/机器人与消费电子/云之间的根本分水岭——比实时性更深，比合规更早。 这一哲学差异决定了汽车 OS 不能简化为 Android、汽车 OTA 不能简化为手机更新、汽车 AI 不能简化为端侧 LLM。

论点四：三类架构原型——垂直闭环、跨设备生态、平台赋能——对应三种不同的”架构控制点”持有策略，没有绝对优劣，只有对资源禀赋、组织能力、监管环境的差异化适配。 Tesla、华为、Google 各自的成功都建立在对自身控制点的清晰认知与极致经营之上。

论点五：AR0—AR5 架构能力门槛框架可以作为跨产业的共同语言，但必须明确”能力门槛而非时间线、低阶不会消失”的方法论原则。 这一原则避免了过度乐观的演进叙事，承认产业的多速发展现实。

论点六：Architecture Intelligence (AI²) 作为新兴研究范式，其方法论与工具的成熟度 (AI²-ML) 是架构决策质量的独立约束维度。 单纯追求被设计系统的架构成熟度而忽视架构智能工具，会导致架构债务的指数级累积。

1.8.2 6.2 七个架构控制点

图 6.1 七大架构控制点网络。控制点持有数量与质量是抵抗”被边缘化”风险的核心指标。

跨案例分析显示，所有成功的架构主导者都至少掌握以下七个控制点中的两到三项：

控制点	含义	典型持有者
C1: 终端入口	拥有海量用户的设备与品牌	Apple、Google (Android)、华为、小米、Tesla
C2: SoC	自研或深度定制核心计算芯片	Apple、Tesla、华为、Google、NVIDIA、高通
C3: OS / 微内核	拥有底层操作系统与内核	Apple、Google、华为、QNX、Microsoft
C4: 中间件	拥有标准化中间件与服务总线	AUTOSAR、ROS 2、Kubernetes、SOAFEE

控制点	含义	典型持有者
C5: 数据闭环	拥有大规模数据采集 → 标注 → 训练的闭环	Tesla、Waymo、华为、 Google、Mobileye
C6: 仿真与世界模型	拥有大规模仿真与生成式数据 能力	NVIDIA、Waymo、Tesla、 Wayve
C7: 安全案例与合规	拥有功能安全/网络安全/AI 安 全的工程方法学与认证能力	Bosch、Mobileye、TÜV、 Exida

最优策略不是盲目全栈，而是识别自己真正可控的”架构控制点”——至少要控制其中两到三项，并在其余领域选择平台合作。这是 CARIAD 案例的核心教训：试图同时控制所有点而无足够内部能力，反而失去全部主导权。

1.8.3 6.3 收敛-分化的终局图景

图 6.3 22 家 OEM 架构成熟度分布 (Snapshot 2026.1.31), 含行业均值线。

会持续收敛的:

- 硅基平台 (Arm Neoverse、Blackwell IP 跨域)
- AI 模型架构 (Transformer + Diffusion + Flow Matching + VLA 跨域通用)
- 异构计算范式与混合关键性 OS
- 软件定义更新机制与持续部署管线
- 边-云 AI 分工与世界模型

会持续分化的:

- 失效哲学与安全壳
- 实时确定性保证
- 责任归属与可审计性
- 法规适配与合规框架
- 物理执行器与传感器形态

终局图景的具象化: 在 AR5 阶段, 仅凭一套底层多核分发微内核 OS + 一个洞悉世界物理定律的基础视觉与逻辑大语言模型, 便能同时丝滑驱动用户的随身全息通讯设备、一辆飞驰在暴风雪山区的无人物理穿梭舱、以及一个正在厨房娴熟颠勺炒菜的双足人形机器人在复杂物理三维世界中优雅协同、自主运作。这是当代各产业架构演进所共同指向的终极愿景——但抵达这一愿景的路径、时间表与具体形态, 仍存在巨大的不确定性。

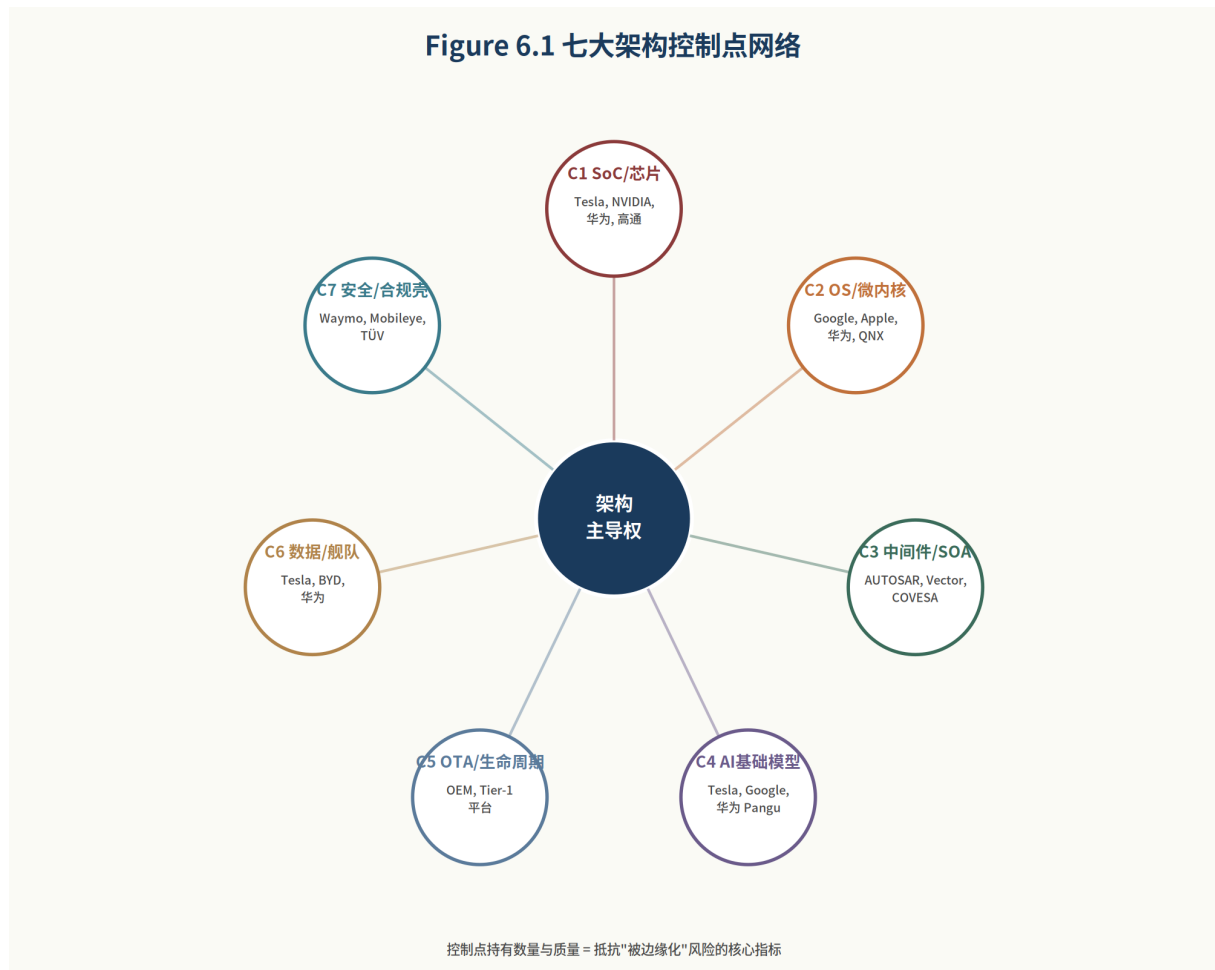


Figure 8: 图 6.1 七大架构控制点网络。控制点持有数量与质量是抵抗“被边缘化”风险的核心指标。

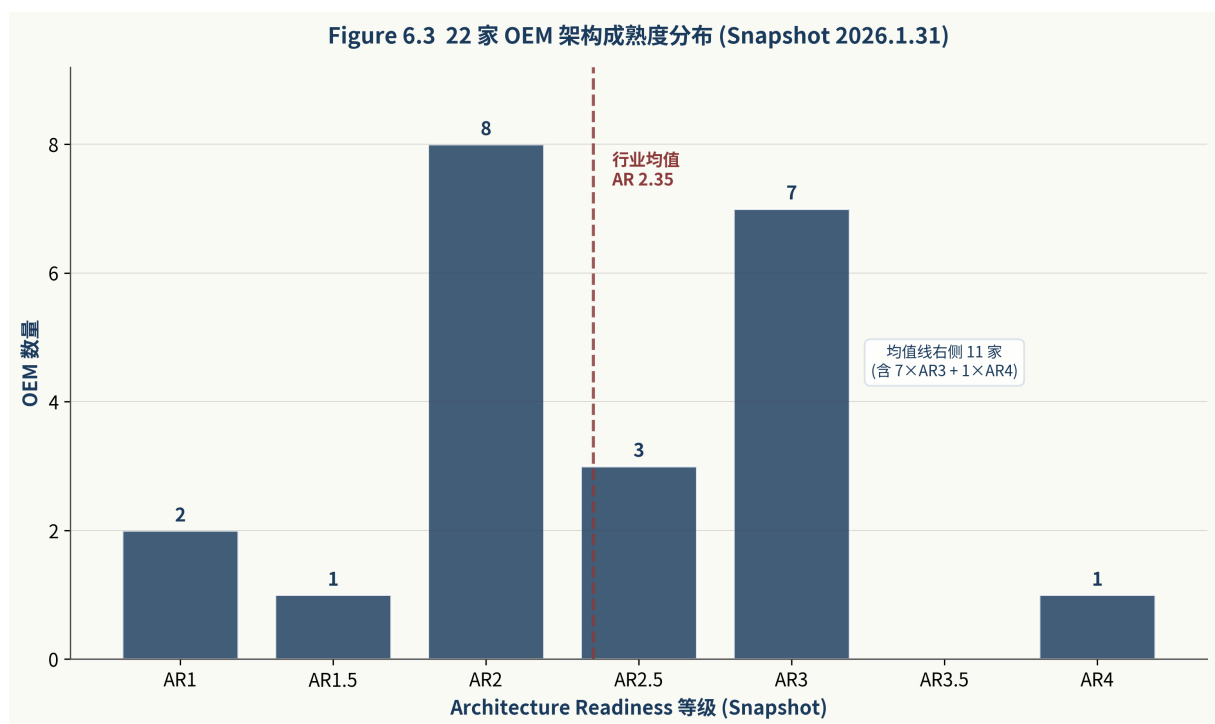


Figure 9: 图 6.3 22 家 OEM 架构成熟度分布 (Snapshot 2026.1.31), 含行业均值线。

1.8.4 6.4 研究边界与限制

本研究作为综述性工作，明确承认以下限制：

限制一：公开信息的不对称。Tesla AI Day、NVIDIA GTC、Google DeepMind 发布、AUTOSAR/AOSP 开源文档提供了大量可验证的技术细节；而华为、小米的车端底层实现公开细节相对有限——许多判断只能基于公开白皮书、年报、产品发布信息进行方向性推断。本文在比较时已力求标注这一不对称性，但读者仍应警惕将“公开度差异”误读为“能力差异”。

限制二：时间外推的不确定性。Tesla AI6/AI7、华为 Pangu 6、Google Gemini 4、Waymo 第 7 代 Driver、Mobileye EyeQ7 等在 2026 年之后的实际落地速度，仍受监管节奏、供应链稳定性、合作伙伴执行、地缘政治多重影响。本文所提及的时间窗口为“基于当前路线图的合理估计”，不构成预测承诺。

限制三：AR0—AR5 框架的研究性质。AR 框架是本研究提出的方法论建议，而非行业共识标准。机器人行业目前尚无等同于 SAE J3016 的全球统一架构成熟度标准。本框架的价值在于提供跨域讨论的共同语言，而非替代任何现有行业标准。框架本身欢迎来自产业与学术社区的批评、修订与扩展。

限制四：AI²-ML 框架的边界。AI²-ML 框架在汽车 E/E 架构、机器人、航天、工业自动化等长生命周期、强约束驱动领域适配度较高；在智能手机、纯互联网应用等迭代快、约束相对宽松的领域适配度较低（智能手机架构更依赖 EDA 工具链而非系统级架构工具）。本框架不主张普适性。

限制五：本文不构成对特定企业的投资建议或产业政策建议。所有案例选取与分析仅服务于方法论研究目的。

1.8.5 6.5 开放问题

本研究在推进过程中识别出以下开放问题，邀请学术与产业社区共同探讨：

1. **端到端神经网络的可证明安全边界在哪里？** 在 EU AI Act 与 ISO/PAS 8800 框架下，纯黑盒端到端系统是否可能获得 L4/L5 责任声明？还是必然需要可解释中间层与运行时监视器？
2. **跨形态泛化的算法极限在哪里？** Tesla 的“占用网络可形态无关”假说能否在更广泛的机器人形态（四足、轮式、飞行）上得到验证？还是会遭遇形态依赖的物理限制？
3. **架构控制点的最小集合是什么？** 是否存在一个最小控制点子集，使得一家公司能够在不掌握 SoC 的前提下仍然主导架构演进？
4. **AR3 到 AR4 的跃迁需要何种组织能力？** 当前产业内仍鲜有完成这一跃迁的范例。CARIAD 困局是孤例还是普遍现象？
5. **AI²-ML 工具的 L4—L5 实现路径？** 用 AI 自动化架构决策面临怎样的形式化、可验证性、信任建立挑战？

1.8.6 6.6 致谢与方法论声明

本研究采用综述性研究方法，材料来源涵盖：

- **学术文献**：IEEE Transactions、SAE Technical Papers、ACM Computing Surveys、Springer Automotive Innovation 等期刊；
- **标准文档**：ISO、SAE、IEEE、UNECE、AUTOSAR、COVESA 等组织发布的标准与白皮书；
- **企业披露**：Tesla AI Day、NVIDIA GTC、Google DeepMind 博客、华为开发者大会（HDC）、小米发布会、Waymo 技术博客、Bosch/Aptiv 白皮书；
- **监管文件**：UNECE WP.29、欧盟委员会、中国工信部 GB 系列标准；
- **行业研究**：McKinsey、Roland Berger、Bain、BCG、Strategy Analytics、IHS / S&P Global Mobility 公开报告；
- **权威媒体**：CnEVPost、CarNewsChina、TechCrunch、The Robot Report、Electrive、Electronic Design 等专业媒体。

本研究在引用具体数字与时间节点时已尽力交叉验证，但因公开信息更新速度差异，部分数据可能与最新事实存在偏差。如读者发现具体事实错误，欢迎反馈以便后续版本更新。

本文采用 **CC-BY 4.0** 知识共享署名 4.0 国际许可协议发布。方法论框架（AR0—AR5、AI²-ML）作为研究贡献向社区开放，欢迎引用、修订与扩展。

本研究由 *archi-intelligence Research Team* 整理。

Architecture Intelligence Research Series · Working Paper 2026-01

—全文完—

1.8.7 6.7 致研究界与产业界的开放邀请

本研究作为综述性工作，明确承认其作为单一研究机构产出的局限性。Architecture Intelligence (AI²) 作为一个新兴研究范式，其方法论成熟、实证检验、跨产业应用都需要更广泛的学术与产业参与。在本研究系列的发布之际，我们向以下几类读者发出具体的合作邀请。

6.7.1 致学术研究界 邀请一：方法论批评与扩展

本研究提出的 AR0-AR5 框架与 AI²-ML 框架是初步的方法论建议，远非定论。我们邀请来自系统工程、软件架构、汽车工程、机器人学、AI 安全等领域的研究者对这两个框架提出批评、修订、扩展。具体可探讨的方向包括：

- AR0-AR5 各阶段的判断指标是否需要进一步形式化？目前的描述偏定性，是否可以引入定量阈值？

- AI²-ML 的 5 维评分维度（架构集中化、软硬解耦、OTA 成熟度、功能安全、算力集中度）是否需要扩展或合并？特别是在机器人与具身 AI 场景中，是否需要新增“形态泛化能力”维度？
- AR 与 SAE J3016 的关系是否需要更精确的形式化？是否存在 AR 与其他成熟度框架（如 CMMI、MLOps Maturity Model）的映射关系？

邀请二：跨产业实证研究

本研究在五大产业（汽车、手机、云、机器人、具身 AI）的对标基于公开信息。我们邀请有产业一线访问权限的研究者开展更深入的实证研究，验证或挑战本研究的核心判断。特别欢迎以下类型的实证工作：

- 单一 OEM 或 Tier-1 内部的 AR 演化案例研究——理想情况下需要 5-10 年的纵向数据
- 跨 OEM 的横向比较研究——基于一致方法论评估多个具体平台
- 失败案例的深度剖析——CARIAD 困局已被本研究讨论，但仍有更多类似案例值得专门研究

邀请三：标准化路径探索

AR 与 AI²-ML 作为研究性框架，其向产业标准的演化需要标准化机构的参与。我们邀请 SAE、IEEE、ISO、INCOSE 等标准组织的相关工作组评估这两个框架的标准化潜力，并欢迎将其作为新的 working group 议题。archi-intelligence Research Team 愿意作为方法论的初始贡献者参与这一过程，但不主张对最终标准的所有权——任何最终标准应由产业共识形成，而非单一研究机构主导。

6.7.2 致 OEM 与 Tier-1 供应商 邀请四：架构自评估与反馈

本研究在 2026-02 报告（《State of E/E Architecture 2026》）中对 22 家 OEM 的参考架构做了 AR 与 AI²-ML 评估。我们承认这一评估完全基于公开来源，难免存在偏差。我们诚挚邀请被评估的 OEM 内部架构团队对本研究的评估结果提供反馈，包括：

- 公开来源遗漏或过时的事实更正
- 架构演进路径的修订（特别是未公开的下一代平台细节）
- 方法论应用中的判断分歧

archi-intelligence Research Team 承诺所有反馈将被认真考虑，重大修正将在后续版本中明确记录。我们也明确声明：接受反馈不意味着放弃独立判断；当反馈与公开证据冲突时，我们保留维持原始评估的权利，但会在文本中标注存在的争议。

反馈渠道：corrections@archi-intelligence.org

邀请五：架构演化的开放讨论

本研究第二章描述的博世六阶段框架已是产业共识，但向 6.0 阶段（车云协同）以及更高阶段（AR4 多体物理 AI 平台、AR5 可信通用智能体）的演化仍存在大量未解问题。我们邀请 OEM 与 Tier-1 的资深架构师就以下问题展开公开讨论：

- 在 AR3 阶段, 跨设备协同的标准化路径是什么? 谁拥有共享身份与上下文的”主权”?
- 在 AR4 阶段, 跨形态算力池化的工程实现路径是什么? 汽车与机器人共享同一基础模型的安全壳如何设计?
- 在 AR5 阶段, “可证明地做”的工程方法学是什么? 现有的功能安全 (ISO 26262) 与 AI 安全 (ISO/PAS 8800) 框架是否足够?

这些问题没有标准答案。我们更看重严肃讨论的过程, 而非仓促结论的产出。archi-intelligence 计划在 2026 年下半年 (ELIV 2026 之后) 组织一系列闭门圆桌讨论, 邀请 OEM 与 Tier-1 的核心架构师参与。有兴趣者请联系: research@archi-intelligence.org

6.7.3 致监管与标准化机构 邀请六: 架构成熟度纳入监管考量

当前的汽车监管框架 (ISO 26262、SOTIF、ISO 21434、UNECE R155/R156) 主要聚焦于具体的功能与组件的安全性。然而, 本研究的核心判断之一是——**架构层面的不可逆性决定了系统全生命周期的安全与合规边界**。一个 AR1 阶段的传统分布式架构与一个 AR3 阶段的跨设备协同架构, 在安全管理上是根本不同的挑战。

我们建议监管机构开始考虑将”架构成熟度”作为新车型监管审查中的一个独立维度。具体可探讨的方向包括:

- 对 AR3 以上阶段的架构, 是否需要新增”跨设备身份与责任管理”的合规要求?
- 对 AR4 以上阶段的架构, 端到端 NN 的可解释性与可审计性是否需要单独的监管框架?
- 对 AR5 阶段 (目前仅在研究中), “可证明地做”的认证体系如何与现有功能安全认证体系融合?

我们承认这些问题的复杂性, 且其答案需要监管机构、产业、学术界的长期合作。archi-intelligence Research Team 不主张特定的监管立场, 但愿意为相关讨论提供方法论支持。

邀请七: 架构智能 (AI²) 作为新独立监管领域

本研究第五章引入的 AI²-ML 框架描述的是**架构决策工具的成熟度**——这是一个目前完全未受监管的领域。然而, 随着 LLM 辅助架构决策、AI 生成架构方案、AI 自动化约束检查等实践的普及, 未来 5-10 年内必然出现以下监管问题:

- AI 辅助生成的架构决策记录 (ADR) 是否具有法律效力?
- 当 AI 工具的架构建议被采纳并导致问题时, 责任如何归属?
- 架构智能工具本身的认证机制 (类似于飞行控制软件的 DO-178C) 是否需要建立?

这些问题目前尚未进入主流监管议程, 但其紧迫性正在快速上升。我们邀请监管机构与法律学者尽早启动相关讨论。

6.7.4 致投资与战略决策者 邀请八：架构维度作为尽调与估值的独立指标

当前的汽车产业投资与战略决策框架（如 Wards Intelligence 的 SDV ScoreCard、PwC Strategy& 的 SDV Maturity Model）主要聚焦于市场份额、销量预测、财务健康度、组织能力等维度。架构层面的判断要么缺失，要么过度简化为“是否使用 zonal 架构”这样的标签。

本研究的核心主张之一是——**架构成熟度（AR）与架构智能成熟度（AI²-ML）应当作为独立的投资与战略决策维度，与传统的市场份额、财务健康度并列考量。**具体而言：

- 一家在 AR4 阶段的 OEM，即使当前市场份额较低，其长期竞争力的天花板可能远高于一家在 AR2 阶段但市场份额较高的 OEM
- 一家拥有 L3-L4 级 AI²-ML 工具的 OEM，其架构演化的“摩擦成本”远低于停留在 L1 级工具的 OEM；这一差异在 5-10 年视角下可能等价于数十亿美元的累积研发效率差距
- Tier-1 供应商在七大架构控制点（C1-C7）中的持有数量与质量，是其抵抗“被边缘化”风险的核心指标

我们邀请投资银行、产业咨询、PE 基金、战略部门将这些维度纳入分析框架。archi-intelligence 计划在 2026 年下半年发布配套的“AR & AI²-ML 投资分析方法论指南”，供有兴趣的机构参考。

6.7.5 致开源社区与开发者 邀请九：参与 archi-intelligence Research Series 的演化

本研究系列采用 CC-BY 4.0 许可发布，允许任何人下载、改编、引用。但更深层的合作邀请是——让本系列成为一个**真正的开源协作研究项目**。具体的合作模式包括：

- 对本研究的方法论与结论提出 issue 与 pull request（GitHub 仓库公开后）
- 贡献新的案例研究——特别是欢迎来自不同地区、不同产业的视角
- 协助翻译——本研究 v1.0 提供中英双语，未来希望扩展到日语、德语、韩语
- 协助可视化——本研究的关键信息图欢迎更优秀的设计师贡献改进版本

我们将在 v1.1 修订版中明确列出所有贡献者（按贡献规模），并视贡献规模考虑给予正式的“contributing researcher”署名。

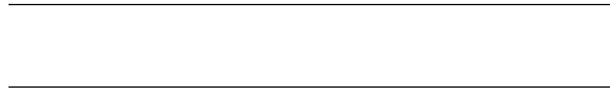
6.7.6 一个研究系列的承诺 最后，我们想以一个具体承诺结束本邀请部分。archi-intelligence Research Team 承诺：

- **持续发布**——本系列计划至少持续 5 年，每年至少发布 2-3 份正式 working paper
- **持续修订**——v1.0 发布后，每 6-12 个月发布修订版，公开记录所有重大修正
- **持续公开**——所有研究方法论、原始数据来源、引用清单 100% 公开
- **持续独立**——拒绝来自任何被评估对象的资金，保持编辑独立性

这些承诺需要时间来检验。我们邀请所有读者持续审视我们的工作,并在必要时给予严厉的批评。一个研究范式的成熟,从来不是单一研究机构的成就,而是整个学术与产业社区的集体产物。

我们期待与您同行。

archi-intelligence Research Team 2026 年 6 月 research@archi-intelligence.org



2 Back Matter 附录

2.0.1 附录 A：术语表 Glossary

按字母顺序排列。每条术语包含中英文对照、简明定义、本报告首次出现章节位置。

A

Architecture / 架构 系统在其环境中的基本概念或属性, 体现在其要素、关系及其设计与演化原则中 (ISO/IEC/IEEE 42010 定义)。本报告中”架构”专指工程语境下的复杂系统组织方式, 不指建筑学意义。→ 首次出现: §1.1

Architecture Description / 架构描述 架构的具体表达形式, 包括但不限于架构图、设计文档、决策记录、约束规范。ISO 42010 强调架构描述比单一架构图更全面。→ 首次出现: §1.3

Architecture Intelligence (AI²) / 架构智能 本研究系列提出的新兴研究范式——应用智能方法 (包括知识图谱、形式化验证、大语言模型、强化学习) 于架构设计、约束检查、参考对标、演化预测的系统性研究。→ 首次出现: §1.7

Architecture Intelligence Maturity Levels (AI²-ML) / 架构智能成熟度等级 本研究系列提出的 5 级评估框架——衡量一个组织在做架构决策时所能调用的方法论、工具、知识库、AI 辅助的成熟度。→ 首次出现: §5.4

Architecture Readiness (AR) Levels / 架构成熟度等级 本研究系列提出的 0-5 级框架——衡量被设计系统 (车辆、机器人、智能终端) 的架构演化阶段。类比 SAE J3016 自动驾驶分级, 但描述的是承载能力的架构成熟度而非自动化能力。→ 首次出现: §5.1

ASIL (Automotive Safety Integrity Level) / 汽车安全完整性等级 ISO 26262 定义的汽车功能安全等级, 分为 ASIL A/B/C/D 四级。ASIL-D 为最高级别, 对应最严格的安全要求。→ 首次出现: §2.1

AUTOSAR (AUTomotive Open System ARchitecture) 全球汽车行业标准化软件框架与 E/E 系统架构。分为 Classic Platform (信号驱动、深嵌入式 ECU) 与 Adaptive Platform (服务化、高算力) 两个并行版本。→ 首次出现: §2.1

C

CARIAD (Car.I.AM Digital) 大众集团软件子公司, 成立于 2020 年。负责大众集团统一软件栈 (Group Software Stack) 的研发。E3 2.0 平台研发滞后是其代表性困局。→ 首次出现: §2.5

Central Compute / 中央计算 E/E 架构 5.0 阶段的核心组件——1-2 个高性能中央计算单元 (HPC) 承担大部分计算工作, 与区域控制器 (ZCU) 协同。→ 首次出现: §2.3

Cross-Embodiment Reuse / 跨形态复用 同一计算栈、感知栈、训练基础设施在不同物理形态 (如汽车与人形机器人) 之间的复用。Tesla FSD 到 Optimus 是典型案例。→ 首次出现: §4.1

E

E/E Architecture / 电子电气架构 车辆内部所有电子硬件、传感器、执行器、计算单元、通信网络与底层软件平台的总体组织方式。→ 首次出现: §2.1

Embodied Intelligence / 具身智能能够在真实物理世界中感知、推理、行动的 AI 系统。本研究中特指人形机器人与跨形态 AI 智能体。→ 首次出现: 引言

F

Fail-Operational / 失效后仍工作汽车与机器人安全哲学——关键功能在部分组件失效后仍能继续工作或安全降级。与消费电子的 fail-soft 形成根本对比。→ 首次出现: §3.7

Foundation Model / 基础模型通过大规模数据训练得到的通用 AI 模型, 可通过微调或提示适配多种下游任务。GROOT、Cosmos、Gemini Robotics 均属此类。→ 首次出现: §3.5

Functional Safety (FuSa) / 功能安全 ISO 26262 定义的汽车安全工程方法学——确保电子电气系统的失效不会导致不可接受的人身伤害风险。→ 首次出现: §2.1

H

HarmonyOS / 鸿蒙操作系统华为自研操作系统, 分为多个版本: HarmonyOS NEXT (5) 于 2024 年 10 月商用, 完全剥离 AOSP 与 Linux 内核; HarmonyOS 6 于 2025 年 11 月发布。→ 首次出现: §4.2.1

HPC (High-Performance Computing) / 高性能计算在 E/E 架构语境中, 指承担车辆主要计算工作的高性能中央计算单元。典型规格: 1,000-4,000 美元单价, 500-2,500 TOPS 算力。→ 首次出现: §2.1

I

ISO/IEC/IEEE 42010 当代架构描述的核心国际标准——将架构定义为”系统在其环境中的基本概念或属性, 体现在其要素、关系及其设计与演化原则中”。→ 首次出现: §1.3

K

Kruchten 4+1 View Model / Kruchten 4+1 视图模型 Philippe Kruchten 1995 年提出的架构描述框架——逻辑视图、过程视图、物理视图、开发视图, 加上整合场景视图。→ 首次出现: §1.3

O

Occupancy Networks / 占用网络 Tesla 2022 AI Day 公开的视觉感知技术——通过体素化 3D 稠密占用 + 占用流场, 从相机生成形态无关的 3D 场景表示。是 FSD 栈可以迁移到 Optimus 的核心技术原因。→ 首次出现: §4.1.2

OTA (Over-the-Air Update) / 空中下载升级通过无线网络对车辆软件进行远程更新的能力。分为非安全关键 OTA (如地图、信息娱乐) 与安全关键 OTA (如 ADAS 控制软件)。→ 首次出现: §2.2

P

Pangu Embodied / 盘古具身智能模型华为云端基础模型 Pangu 在具身机器人场景的专化版本。被应用于华为 Kuafu 人形机器人。→ 首次出现: §4.2.1

S

SDV (Software-Defined Vehicle) / 软件定义汽车以更少、更强的车载计算节点取代大量孤立控制器, 并通过软件在全生命周期持续升级的汽车形态。→ 首次出现: §2.1

SOA (Service-Oriented Architecture) / 面向服务的架构软件架构范式——将系统

功能组织为可独立部署、独立演化的服务。在 SDV 语境中通常通过 SOME/IP 或 DDS 等中间件实现。→ 首次出现: §2.1

System 1 / System 2 在具身 AI 架构中借用 Daniel Kahneman 的概念——System 1 指快速、直觉的感知与反应; System 2 指慢速、深思熟虑的推理与规划。Waymo Foundation Model 等架构采用了端云双模型协同的实现。→ 首次出现: §3.5

T

TOGAF (The Open Group Architecture Framework) 企业架构领域的核心方法学。将架构定义为”系统的组件结构、它们之间的相互关系,以及指导其设计和长期演进的原则与规范”。→ 首次出现: §1.3

Trade-off Decision / 权衡决策 本研究强调的架构核心活动——在功能性需求与非功能性需求之间显式化、量化、可追溯地做出取舍。→ 首次出现: §1.4

V

Vehicle-to-Cloud / 车云协同 E/E 架构 6.0 阶段的核心特征——算力在车端与云端动态分配;端到端模型与世界模型协同;持续舰队学习。→ 首次出现: §2.3

Vitruvius / 维特鲁威 公元前 30-15 年罗马建筑师,著有《建筑十书》。提出”维特鲁威三原则”(Firmitas / Utilitas / Venustas),是架构作为权衡艺术的最早形式化表达。→ 首次出现: §1.1

VLA (Vision-Language-Action) Model / 视觉-语言-动作模型端到端 AI 模型——将视觉感知、语言推理、动作输出统一在单一神经网络中。理想 Mind GPT VLA、Waymo Foundation Model 等均属此类。→ 首次出现: §3.5

W

WEWA (World Engine + World Action Model) 华为 ADS 4.0 架构——云端 World Engine 生成 1,000 倍极端场景密度 + 车端 World Action Model 原生驾驶基础模型。→ 首次出现: §2.4

Why > How (Why Dominates How) / Why 凌驾于 How 本研究强调的架构第二定律——架构的战略价值在于”为何如此选择”的论证逻辑,而非”如何实现”的技术选型。→ 首次出现: §1.4

Z

Zonal Architecture / 区域化架构 E/E 架构 5.0 阶段的核心组件——以物理位置(前左、前右、中、后左、后右等区域)划分的控制器,通过高速骨干网络连接至中央计算。代表是 Tesla HW4。→ 首次出现: §2.3

Zone Control Unit (ZCU) / 区域控制器区域化架构的核心组件——典型规格: 50-70 美元单价,负责特定物理区域的传感器接入、本地控制、电源分配。→ 首次出现: §2.1

2.0.2 附录 B: AR & AI²-ML 评分快速参考卡

本快速参考卡为产业从业者提供 AR & AI²-ML 评分的简化判断工具。完整方法论参见 §5.1-5.5 及 archi-intelligence Research Series 2026-02。

PART A: AR0-AR5 快速判断流程

步骤 1: 判断系统的”被设计实体”是什么

- 汽车 (含 ICE、HEV、BEV、PHEV)
- 机器人 (工业、协作、服务、人形)
- 智能终端 (手机、平板、可穿戴、智能家居)
- 其他复杂系统

步骤 2: 从下表向下匹配, 找到第一个不满足的条件

是 AR0 吗? - ✓ 一项功能对应一个独立控制器 - ✓ 软硬件深度耦合 - ✓ 几乎无 OTA 能力 - ✓ ECU/控制器数量 > 30 (汽车) 或机械臂等单体系统 - → 若全部满足 → AR0 - → 若有不满足 → 检查 AR1

是 AR1 吗? - ✓ 5 ± 2 个功能域控制器 - ✓ AUTOSAR Classic 或 Adaptive 部分采用 - ✓ OTA 已规模化但仍以单 ECU 为单位 - ✓ 跨域通信受物理总线带宽限制 - → 若全部满足 → AR1 - → 若有不满足 → 检查 AR2

是 AR2 吗? - ✓ 区域控制 + 中央计算 - ✓ 服务化诊断与统一信号目录 - ✓ Hypervisor 虚拟化已部署 - ✓ 控制器数量典型 7-15 个 - ✓ 整车软件平台化, 跨车型复用 - → 若全部满足 → AR2 - → 若有不满足 → 检查 AR3

是 AR3 吗? - ✓ 车/手机/穿戴/家居/机器人共享身份与上下文 - ✓ 跨设备任务连续性 - ✓ 边云协同调度 - ✓ 跨形态算力池化 (初步) - → 若全部满足 → AR3 - → 若有不满足 → 检查 AR4

是 AR4 吗? - ✓ 训练、仿真、部署闭环统一 - ✓ 世界模型与基础模型跨车/机器人共享 - ✓ 模型复用率高 - ✓ 策略迁移效率高 (跨形态泛化) - → 若全部满足 → AR4 - → 若有不满足 → 检查 AR5

是 AR5 吗? - ✓ 通用策略层可跨多种实体执行 - ✓ 持续证明安全/合规/责任 - ✓ “可证明地做”已工程化实现 - → 若全部满足 → AR5 - → 若有不满足 → 仍在 AR4 或更低

PART B: AI²-ML 5 级快速判断流程

被评估对象: 架构工具/方法论/组织能力

是 L0 吗? - ✓ 使用 Visio、Miro、PowerPoint 等通用绘图工具 - ✓ 架构图无形式化语义 - ✓ 改动一处其他不会自动更新 - → 若全部满足 → L0 Drawing

是 L1 吗? - ✓ 使用 PREEvision、Capella、Cameo 等专业建模工具 - ✓ 架构有形式化语义 - ✓ 但缺乏跨元素的约束检查 - → 若全部满足 → L1 Modeling

是 L2 吗? - ✓ 模型支持约束规则与一致性检查 - ✓ 可发现 ASIL 违反、带宽超限等具体问题 - ✓ 工具能自动验证设计的内部一致性 - → 若全部满足 → L2 Constraint-Checked

是 L3 吗? - ✓ 工具内置参考架构库 - ✓ 能基于行业对标进行架构评估 - ✓ 支持决策历史的结构化记录 (ADR) - → 若全部满足 → L3 Reference-Aware

是 L4 吗? - ✓ 工具用 AI 辅助架构决策 - ✓ 支持自然语言交互 - ✓ 自动布局、约束推理、trade-off 量化 - → 若全部满足 → L4 AI-Assisted

是 L5 吗? - ✓ 工具能在给定目标与约束下自动生成架构 - ✓ 自主演化架构以应对环境变化 - ✓ 当前尚无成熟产品 → 谨慎判断 - → 若全部满足 → L5 Autonomous

PART C: AR × AI²-ML 二维定位的实战使用

在二维坐标系中, 横轴 L0-L5, 纵轴 AR0-AR5。

- “理想区”位置: 横纵相当 (即工具成熟度匹配系统成熟度)
- “债务区”位置: 纵轴高于横轴 (系统先进但工具落后)
- “过度投资区”: 横轴高于纵轴 (工具先进但系统落后)

典型危险信号:

- 目标 AR3 但仅有 L1 工具 → 高架构债务风险
- 目标 AR4 但仅有 L2 工具 → 极高架构债务风险
- 工具领先系统 3 级以上 → 投资回报不明

典型健康信号:

- 工具与系统成熟度差距 ≤ 1 级
- 工具与系统协同演化 (每提升一级, 另一个也跟进)

2.0.3 附录 C：跨产业 5 维对标完整数据表

本附录提供第三章跨产业基准的完整原始数据。所有数据来源于公开发布的官方信息、上市公司财报、专利文献、行业标准组织发布、主流技术媒体。

DIMENSION 1: HARDWARE / 硬件

汽车 E/E 主要 SoC (2025-2026 量产/即将量产):

- Tesla AI4 (HW4): 三星 5nm, ~500 TOPS INT8
- Tesla AI5: TSMC N3P, 2,000-2,500 TOPS (2026 末/2027 初)
- Tesla AI6: 三星 Texas, 训练 + 推理融合 (开发中)
- NVIDIA Thor: 7nm, 2,000 TFLOPS FP8, 77B 晶体管
- 高通 SA8775P: 5nm, 舱驾融合, ASIL-D
- 华为 MDC 1000: 基于 Ascend 910B, 1,000 TOPS
- 蔚来神玑 NX9031: 5nm, 50B+ 晶体管, ~2,000 TOPS 等效
- 小鹏 Turing: 40 核 + 双 NPU, 750 TOPS/单芯
- Mobileye EyeQ Ultra: 12× 双线程 RISC-V + 64 加速核, <100W

智能手机/消费电子 SoC (2025):

- Apple A19: 3nm, Neural Engine ~45 TOPS
- Apple M5: 3nm, 桌面级移动 SoC
- 高通 Snapdragon 8 Elite Gen 5: 3nm, AI 优化
- 联发科 Dimensity 9500: 3nm
- 华为 Kirin 9030: SMIC N+3 (5nm 级), 2025 量产
- Google Tensor G5: TSMC 3nm, 2025 年 8 月 (脱离三星代工)
- 小米 Xring O1: 3nm, 首款应用处理器

云端 AI 训练芯片 (2025-2026):

- NVIDIA GB200 NVL72: 1.44 EFLOPS FP4
- Google TPU v7 Ironwood: 4,614 TFLOPS FP8 单芯, 9,216 芯 superpod
- AWS Trainium2: 4.6 PFLOPS
- 华为 CloudMatrix 384: 300 PFLOPS BF16, 384 颗 Ascend 910C 光互联

机器人芯片 (2025-2026):

- NVIDIA Jetson Thor: 2,070 TFLOPS FP4, 机器人专用
- Tesla AI4 (复用): 用于 Optimus Bot Brain
- 华为 Ascend (车规版): 用于 Kuafu 机器人
- 特定关节执行器芯片: 多家供应商专有方案

DIMENSION 2: SOFTWARE / 软件

汽车 OS / 中间件 (2025-2026 主流配置):

- AUTOSAR Classic: R24-11 版本, 信号驱动, 嵌入式 ECU
- AUTOSAR Adaptive: R24-11 版本, SOA, POSIX, 高算力
- QNX: 黑莓 QNX 8.0 (2025), 实时 RTOS
- Linux Automotive: Yocto / AGL 主流, BSP 厂商定制
- Tesla Auto OS: Tesla 自研, 不公开细节
- HarmonyOS-Auto: 华为车机版, 与 HarmonyOS NEXT 同源
- 鸿蒙智行 HiCar: 华为车手互联

智能手机/消费电子 OS (2025-2026):

- iOS / Darwin: 18.x, Apple 独占
- Android (AOSP): 15/16, Google 主导
- HarmonyOS NEXT: 5/6, 微内核, 完全剥离 AOSP/Linux
- HyperOS: 1.0/2.0/3.0, Linux + Vela 双内核

云端 OS / 编排:

- Linux + Kubernetes: 2025 年生产采用率 82%
- 服务网格 (Istio/Linkerd): 标准微服务通信层
- Serverless: AWS Lambda / GCP Cloud Functions

机器人栈:

- ROS 2: Jazzy Jalisco (LTS, 2024 年 5 月)
- DDS: ROS 2 通信中间件
- Nav2: 机器人导航栈
- MoveIt 2: 运动规划栈
- Apex.AI: ASIL-D 认证 ROS 2

具身 AI 模型架构:

- VLA (Vision-Language-Action):
 - Tesla FSD V12-V14
 - 理想 Mind GPT VLA
 - 华为 ADS 4.0 WEWA
- Diffusion Policy: 研究主流
- Flow Matching: 新兴方向

- System 1 / System 2: Waymo Foundation Model 等

DIMENSION 3: DATA / AI

汽车 AI 模型与训练规模:

- Tesla FSD V13: HW4 原生分辨率, 4.2× V12 数据
- Tesla FSD V14: 4.5× V13 参数 (2025-26)
- 华为 Pangu 5.5: 718B 参数
- 华为 ADS 4.0: 87.6 亿公里累计辅助驾驶里程 (2026.1)
- 理想 VLA: 12 亿公里数据训练
- 比亚迪 DiPilot: 日均 7,200 万公里数据 (44 万 + 车辆舰队)
- 小鹏 Turing 端到端: 多模态训练

智能手机端侧 AI:

- Apple Intelligence: 3-8B 模型本地 + 私有云计算 (PCC)
- Google Tensor: 3-8B 模型 + Cloud TPU
- 华为 Pangu (端侧): LoRA 微调适配

云端基础模型:

- OpenAI GPT-4o / o1: 具体规格不公开
- Anthropic Claude: 具体规格不公开
- Google Gemini 3: 2025 年发布, 完全在 TPU 训练
- 华为 Pangu 5.5: 718B 深思考 + 5 个行业子模型
- DeepSeek V3 / R1: 公开论文

机器人基础模型:

- NVIDIA GR00T N1: 2025 GTC 发布
- NVIDIA Cosmos: 世界模型, 跨车-机器人共享
- Physical Intelligence $\pi 0 / \pi 0.5$: 通用机器人模型
- Figure AI Helix: 端到端 VLA
- Google Gemini Robotics: 跨形态泛化
- Skild AI Skild Brain: 通用控制基础模型
- OpenVLA (开源): 7B VLA 模型

DIMENSION 4: ECOSYSTEM / 生态

汽车生态模式:

- OEM 全栈 (Tesla): 硅-OS-应用-云-机器人单一企业全栈

- Tier-1 联盟 (华为 HIMA): 鸿蒙智行五界 (问界/智界/享界/尊界/尚界)
- 平台 (NVIDIA/高通): 芯片 + 中间件 + 开发工具给多 OEM

消费电子生态模式:

- 围墙花园 (Apple): iOS 独占, App Store 30% 抽成
- 平台共享 (Android): AOSP 开源, GMS 服务商业化
- 垂直整合 (华为): HarmonyOS NEXT 全自研, 元服务生态

云生态模式:

- 开源主导: CNCF 15.6M 开发者
- 超大规模商用: AWS / Azure / GCP / 阿里云 / 腾讯云

机器人生态模式:

- 垂直整合: Tesla Optimus, Figure AI
- 水平平台: NVIDIA Isaac, Skild AI

具身 AI 生态模式:

- NVIDIA 三计算机范式: 训练 (DGX) - 仿真 (Omniverse) - 运行 (Jetson Thor)
- 跨形态共享: Cosmos 世界模型已被 Toyota/1X/Skild/Figure/Agility/Uber/Waabi/Foretellix/小鹏/Galbot/Fourier 同时采用

DIMENSION 5: SAFETY & COMPLIANCE / 安全合规

汽车安全合规标准:

- ISO 26262: 功能安全 (FuSa), 2018 版
- SOTIF (ISO 21448): 预期功能安全, 2022 版
- ISO 21434: 网络安全, 2021 版
- UNECE R155/R156: 网络安全与软件更新法规
- ISO/PAS 8800: AI 安全 (2024 发布)

智能手机/消费电子合规:

- TEE / Secure Enclave: 硬件级安全飞地
- GDPR / PIPL / CCPA: 数据隐私法规
- E2E 加密: 端到端加密
- 注意: 无强制 FuSa 要求

云端合规:

- SOC 2: SaaS 安全审计
- ISO 27001: 信息安全管理
- 零信任 (Zero Trust): 网络架构范式
- 机密计算: SGX / TDX / Nitro Enclaves

机器人合规:

- ISO 13482: 个人护理机器人, 2014 (2025 更新)
- ISO 10218: 工业机器人, 2025 更新
- ISO/TS 15066: 协作机器人
- ISO 13849: 机械安全

具身 AI 合规:

- 标准空白: 目前无全球统一标准
- 新提议:
 - DeepMind ASIMOV Benchmark (2024)
 - Anthropic Constitutional AI (扩展到机器人)
 - 各国监管讨论中 (EU AI Act 相关延伸)

2.0.4 附录 D: archi-intelligence Research Series 简介

关于本系列

archi-intelligence Research Series 是 archi-intelligence 研究机构定期发布的跨产业架构研究综述系列。本系列以建立 Architecture Intelligence (架构智能, AI²) 作为独立研究范式为目标, 通过严格的方法论、公开的数据来源、独立的编辑判断, 为产业界与学术界提供可信、可引用、可演化的研究资产。

发布频率与计划

本系列计划至少持续 5 年 (2026-2030), 每年发布 2-4 份正式 working paper。已规划的发布如下:

2026 年发布计划:

- 2026-01 架构的世纪迁徙: 从工程本体论到具身智能的统一图景 (旗舰主报告)
- 2026-02 State of E/E Architecture 2026: 22 家 OEM 的 AR & AI²-ML 评估
- 2026-03 Tesla FSD-Optimus Unified Stack 解剖: 一个 AR4 垂直闭环案例的深度剖析
- 2026-04 TBD (Year 2 启动)

2027-2030 年规划方向 (具体题目待定):

- 中国 E/E 架构深度专题
- 机器人 EEA 浮现: 从原型到平台
- AI² Foundation 治理章程白皮书
- 跨产业架构标准化路径
- 失败案例集 (CARIAD 之外的深度剖析)

每份 working paper 的版本演化

每份 working paper 采用语义化版本管理:

- v1.0 - 首次正式发布
- v1.x - 重大修订 (通常每 6-12 个月一次)
- v2.0 - 方法论或核心论点的根本性更新

每次版本更新会公开记录变更日志 (Changelog), 明确:

- 修正的事实错误
- 调整的方法论
- 新增的案例与数据
- 改变的核心论点
- 致谢的反馈贡献者

引用与传播

本系列所有 working paper 在 CC-BY 4.0 许可下发布, 欢迎:

- 学术引用 (请使用每篇报告的标准引用格式)
- 媒体报道 (请保留 archi-intelligence 出处标注)
- 教学使用 (无需事先授权)
- 翻译传播 (v1.0 提供中英双语, 欢迎社区贡献其他语言)
- 二次创作与衍生研究 (请明确标注与原作的差异)

不允许:

- 删除或修改作者归属
- 将本作品作为商业付费产品销售 (CC-BY 不限制商业使用, 但要求保留许可与归属, 且不能虚构对作品的独占权)
- 利用本作品做虚假背书或诱导性宣传

archi-intelligence 机构信息

archi-intelligence 是致力于 Architecture Intelligence 研究范式建立的独立学术研究机构。本机构在 v1.0 发布阶段由 Arkimind (一家专注于汽车电子电气架构智能化的商业实体) 提供 seed sponsorship。详细的治理结构、利益冲突管理、编辑独立性保障措施请参阅本报告 COI Disclosure 与 archi-intelligence.org/governance。

我们计划在 Year 3 (2028 年) 将研究业务独立为 AI² Foundation 非营利机构, 进一步强化与商业实体的隔离。这一演化路径完全公开, 欢迎所有读者持续审视。

如何参与

- 研究合作 / 同行评审: research@archi-intelligence.org
- 内容更正 / 反馈: corrections@archi-intelligence.org
- 媒体咨询: press@archi-intelligence.org
- 赞助洽谈 (限非被评方) : sponsors@archi-intelligence.org
- GitHub: <https://github.com/archi-intelligence>
- LinkedIn: <https://linkedin.com/company/archi-intelligence>
- 官网: <https://archi-intelligence.org>
- DOI: <https://zenodo.org/communities/archi-intelligence>

最后致谢

本系列的每一份 working paper 都是无数前人工作的延续。我们对所有在公开来源中分享技术细节、研究方法、案例数据的研究者、工程师、记者表达深切感谢。一个研究范式的成熟, 从来不是单一研究机构的成就, 而是整个学术与产业社区的集体产物。

We stand on the shoulders of giants. So shall those who follow us.
archi-intelligence Research Team 2026 年 6 月



3 编辑修订指南

D1 旗舰主报告—整体编辑修订指南

本指南为 LaTeX 排版与最终校对阶段提供具体编辑标准。

A. 术语统一

以下术语全文统一使用, 禁止变体:

- ✓ Architecture Intelligence (首次出现完整词)
- ✓ AI² (后续简称, 使用上标²)
- ✓ AI²-ML (架构智能成熟度等级)
- ✓ AR0, AR1, AR2, AR3, AR4, AR5 (无空格)
- ✓ E/E Architecture (注意斜杠)
- ✓ SDV (软件定义汽车, 避免与 SUV 混淆)
- ✓ ISO 26262 (空格分隔, 不写 ISO26262)

避免变体:

- ✗ AI2 (易与 Allen Institute AI2 混淆)
- ✗ AI 平方 / Architectural Intelligence
- ✗ A.R. / AR Level (统一为 AR0-AR5)

B. 句式风格

学术综述风格基调:

- 客观陈述, 避免主观情绪词 (“显然”、“无疑”、“令人震惊”)
- 区分事实陈述与判断陈述
- 使用 “本研究认为”、“本研究的核心判断是” 表达原创观点
- 使用 “据公开报道”、“基于公开来源推断” 表达不确定信息

句长控制:

- 平均句长 25-40 字 (中文) / 20-30 词 (英文)
- 超长句拆分 (中文 > 60 字 / 英文 > 40 词需要拆分)

C. 数据标注规范

所有具体数字必须标注来源或时间:

- ✓ “Tesla AI4 (HW4): 约 500 TOPS (Tesla AI Day 2023)”
- ✓ “蔚来神玑 NX9031: 50+ 亿晶体管 (2024.7 NIO IN 发布)”

模糊数字使用范围或限定词:

- ✓ “数百亿美元投入”
- ✓ “约 2-3 周工作量”

避免：

- ✗ “据说有 1000 TOPS”（来源不明）
- ✗ “可能在 2027 年量产”（无依据的预测）

D. 时态规范

中文写作：

- 已发生事件用过去时式（“2024 年 10 月发布”）
- 正在进行用现在时式（“目前正处于演化中”）
- 未来事件用条件式（“计划于 2027 年量产”）

英文写作：

- 已发生事件用 past tense
- 描述当前状态用 present tense
- 未来计划用 “is planned to” / “is expected to”
- 避免用 “will” 做强预测

E. 关键陈述的修辞控制

第六章 6.3 节“终局图景”招牌段保留原始修辞张力, 但其他章节避免过度修辞。具体原则：允许的修辞工具：

- 简短的隐喻（“数字石器时代”、“硅基生命体”）
- 历史类比（“维特鲁威三原则” → “可靠性—功能性—可维护性”）
- 数据对比（“30-40% 年增速 vs 1% 年化收缩”）

避免的修辞：

- 戏剧化形容词（“令人震惊”、“颠覆性”用于描述常规演进）
- 价值判断（“先进”、“落后”，除非有客观指标支持）
- 单边立场（避免“中国领先”或“欧美落后”的简单结论）

F. 图表引用规范

正文中引用图表的标准格式：

- 中文：“如图 4.1 所示” / “见表 2.1”
- 英文：“as shown in Figure 4.1” / “see Table 2.1”

图表编号规则：

- 图表 X.Y,X 为章节号,Y 为顺序号
- 跨章节引用使用完整编号 (“详见图 5.2”)

G. 参考文献规范

引用格式: APA 第 7 版

期刊文章: Author, A. A. (Year). Title of article. Journal Name, Volume(Issue), pages. <https://doi.org/xxxxx>

会议论文: Author, A. A. (Year). Title of paper. In Proceedings of Conference Name (pp. xxx-xxx). Publisher.

公司技术发布: Tesla, Inc. (2023, May 18). FSD V12 release notes. <https://www.tesla.com/>
...

标准文件: International Organization for Standardization. (2018). Road vehicles —Functional safety (ISO 26262:2018).

对所有具体数据点, 必须有可追溯的来源。来源包括但不限于:

- 上市公司财报 (年度 10-K, 季度 10-Q, 中报)
- SEC / HKEX / SSE 公开文件
- 公司官方技术发布与新闻稿
- 已公开的专利文献
- 同行评审学术论文
- 行业标准组织官方发布
- 主流技术媒体的署名报道

不接受的来源:

- 匿名爆料 / Twitter 传闻
- 未署名的中文公众号转载
- 缺乏可追溯链接的二手引用

H. 双语版本同步规范

中英双语版本采用”双源同步”而非”翻译”模式:

中文版与英文版的内容差异允许范围:

- ✓ 案例引用的本地化 (中文版多引中文资料, 反之亦然)
- ✓ 表达习惯差异 (中文重意合, 英文重形合)
- ✗ 核心论点的差异
- ✗ 数据与结论的不一致
- ✗ 图表编号与内容的不一致

每次重大修订后, 双语版本需做交叉校对, 确保:

- 章节顺序一致
- 论点表达一致
- 数据完全一致
- 图表完全一致

I. 法律与合规复核要点

发布前必须法律复核的章节：

- 第二章 §2.5 (CARIAD 困局)
- 第四章 §4.1 (Tesla)
- 第四章 §4.2 (华为/小米)
- 第四章 §4.3 (Google/Waymo)
- 第六章 §6.3 (终局图景)

复核流程：

1. archi-intelligence Research Team 内部审阅
2. 独立德国劳动律师就 VW 竞业条款 sign-off
3. 独立产业法律顾问就潜在 OEM 反应 sign-off
4. Advisory Board 5 位前 OEM 架构师同行评审
5. 最终主编批准发布

J. 视觉信息图制作交付要求

每张关键信息图需提交以下格式：

- 高分辨率 PNG (300 DPI, 用于打印)
- SVG 矢量格式 (用于网页与缩放)
- LaTeX 兼容的 PDF (用于嵌入主文档)
- 中英双语文字版本 (图本身共用, 标签文字双语化)

设计师交付清单 (共 13 张关键图) :

Figure 1.1, 1.2, 1.3 Figure 2.1, 2.2, 2.3 Figure 3.1, 3.2 Figure 4.1, 4.2, 4.3, 4.4,
4.5 Figure 5.1, 5.2 Figure 6.1, 6.2

设计风格统一标准：

- 单色为主 (archi-intelligence 深蓝 #1a3a5c)
- 强调用暗红 #8b3a3a (仅用于关键论点)
- 字体: Inter / 思源黑体
- 留白充足, 信息密度适中
- 参考: Anthropic 论文风格 / MIT Press 出版物

最终发布前 Checklist (必须 100% 完成)

- 所有章节内容定稿
- 所有数据已二次核对
- 所有图表已交付
- 中英双语版本已交叉校对
- DOI 已申请并激活
- Zenodo 上传已完成
- GitHub 仓库已设置
- archi-intelligence.org 发布页面已搭建
- Press Kit 已准备 (中英双语)
- LinkedIn 系列文章已草拟
- 媒体 embargo 列表已确认
- 法律 sign-off 已获得
- Advisory Board 同行评审已完成
- 中文版与九章智驾合作已协调
- ELIV 2026 现场分发版印刷已下单

完成以上 Checklist 后, 进入正式发布。
