

新上汽 新技术 新动能

开栏语

随着新发展理念深入人心，上汽各企业坚持创新驱动，努力提升核心竞争力。从本期开始，本报开设“新上汽，新技术，新动能”专栏，重点关注上汽在各个创新领域里的技术发展成果。

平行世界里，上汽起建汽车工厂

本报记者 顾行成

“在信息科技时代，数据就是一种生产资料。各种工具软件、工具包、AI工具链，就像是一台台在云端上的数字化机器人，对数据进行分类、筛选、分析、标注、训练。在各种深加工之后，再根据需求，产出各种数据产品，包括软件、算法模

型，或是写入车载芯片里的一段程序。我们把现在正在做的这个项目，称之为‘数据工厂’。”

上海帆一尚行科技有限公司（以下简称“帆一尚行”）AI产品部总监谭黎明用一段形象的比喻，揭开了不久前刚在TECH SHOW上露面的“数据工厂”的神秘面纱。



1 平行世界里的汽车工厂

不少工程师都喜欢把这个专注于自动驾驶研发的“数据工厂”，称之为在平行世界里的数字化研发汽车工厂。

经过两年多的发展与建设，帆一尚行已经在这个数据工厂里建成了集数据导入、数据标注、数据分析处理、模型训练、仿真检验等为一体的全流程闭环工具链，具备了完整的数据运营功能，并为洋山港智能重卡、智己汽车、R标智驾、商用车双J3项目和Robotaxi等上汽集团重点战略项目的实施提供了有力的支撑。

帆一尚行总经理钟中告诉记者：“十四五”期间，帆一尚行将致力于把AI和数据技术赋能更多的汽车设计、生产、制造、营销和运营场景的应用，充分发挥自主研发的优势，服务好各个重大智能驾驶研发项目。

目前，我国的产业结构正在发生调整，热点和主线正从第三产业向先进智能制造转变。中国互联网已进入“下半场”，即从消费互联网转变为产业互联网，将互联网的创新应用成果

深化于各项领域，支撑经济社会数字转型、智能升级和融合创新。

钟中表示，帆一尚行正在打造的行业云——帆一云，专注于新基建、智能制造、人工智能、智能网联和网络安全领域，已经成为面向政企客户输出领先技术、标准化产品、非标产品和增值服务的数智化云平台。

据了解，目前这个“建设”中的数据工厂，已经实现了数据的分层、互通和共用，这使得上汽集团各研发项目实现了数据资产的价值最大化，同时缩短了模型训练时间、拓展了模型训练场景、大幅度降低了研发成本，产品迭代速度也得到了提升。这些都成为上汽集团战略目标的有效支撑力，并开拓了新一代智驾数字化基础运营平台。

“建设过程中困难很多。”谭黎明坦言，“每个客户对数据的诉求不一样、标准不一样，安全规范也不一样。我们在开始决定做这个数据工厂的时候，甚至没有一个可参考学习的对象。”

2 从零开始铸就灵魂

2018年6月，帆一尚行的工程师写下了数据工厂的第一行代码。

“其实，在此之前，整个行业内都没有‘数据工厂’这个概念，只有一些零散的工具链。例如，数据的筛选、标注、训练，以及各种场景下的单点功能，包括面部识别、自动泊车等。”谭黎明告诉记者，“不同企业之间各自的标准和定义不同，各个工具链之间没有形成整合，哪怕我手握再多的工具链，它们也只是一条条单独的生产线，还不能说是一个完整工厂的概念。”

整个行业都在群雄逐鹿，显然并不存在一个整合好的大数据工厂。团队在2019年年年初毅然决定，将现有架构彻底重构。

“那段时期，上汽集团内部关于智能驾驶的需求越来越多。如果还是一个工具链‘各自为营’，就无法满足整个集团的智能驾驶需求。从那个时候开始我们就想，干脆从零开始，建立一个数据工厂，把所有工具链都整

合起来，还能把核心技术都掌握在自己手中，为未来竞争占据制高点。”谭黎明说道。

以上汽云计算中心为支撑、研发迭代为载体，广泛吸收头部厂商的先进经验，并结合上汽自主研发的自动驾驶技术，数据工厂团队在两年多的时间内，自主研发、融合打造了各类模块30余个，已基本上满足全场景应用，并同步完成了各类数据生产线的统筹优化。数据工厂的建设进度也在快速推进中，基本上每两周迭代一次小版本，每半年更新一次大版本。截至目前，团队已在现有工厂的雏形上累计筛选出10PB的价值信息（占信息总量的20%），并完成了两次大规模的工厂重构。

“再强的算法，再聪明的人工智能，想要发挥作用，首先要把数据存起来、标注好，并且能够用起来，而我们就是负责打造一套这样的基础设施。它和物理工厂最大的不同之处在于，我们这个建在云上的工厂，它的产品，也

就是数据是可以被复制利用的。”谭黎明自信地告诉记者，“我们现在的这个数据工厂是把各种

工具链相互打通，实现互用。当然，数据的归属及权限管辖，我们是做严格控制的。”

3 从云端走下去“降维”打击

不久前，数据工厂和郑州相关部门签署了合作协议，共同发展当地智能网联产业。据了解，数据工厂将一部分功能模块和工具链挑选出来，根据当地的需求，组合成一个全新产品并对其进行部署。例如，帆一尚行和郑州工信局旗下的一家公司进行合作时，对方就需要一个云控平台，用于分析进入指定示范区的车辆上传到云上的数据。在一个固定场景内，通过数据工厂输出模块，训练当地的交通工具，实现有限的自动驾驶，同时将数据传回云端。落脚点于具体的应用场景，而不是针对某个车企，这是帆一尚行在面对数据工厂未来如何打开市场的一计妙招。这样不仅可以避免一些敏感问题，还能和目标形成有效互补。记者在采访过程中了解到，诸如此类先在工厂中选取部分工具链进行

定制整合，再运用到合适的场景应用模式，被帆一尚行的工程师们形象地称为智驾技术中的“降维打击”。

除了拓展应用场景，数据工厂还能通过不同工具链的组合，形成不同级别的自动驾驶产品包。比如，当前占比最大的低级别自动驾驶量产车，在未来也可以通过数据工厂，用更高级的自动驾驶技术去降维反哺，进行兼容覆盖。除此以外，像无人送餐车、环卫车等有限固定场景中的载具，也都是谭黎明团队“盯”上的目标。

记者离开时，正好看到有几名工程师不顾烈日当空，正埋头围着一辆无人送餐车不停地测试。“我们数据工厂下线的产品，一开始也只能打60分。”谭黎明谦虚地表示，“另外40分，得到场景中获取。闭门是造不好车的，云上、云下都一样。”