



## 算法争先 AI助力基因检测

人工智能抗击新冠肺炎疫情，多以测温、无人配送的形式冲锋陷阵，就算是CT智能诊断等临床方向，也鲜有像阿里达摩院算法专家顾斐这样去疾控中心一线支持基因检测的经历，从1月28日开始，顾斐已经在疾控中心奋战了一个多月。这是顾斐首次现场支持合作伙伴。因为他和团队把病毒检测速度由数小时缩短到半小时，病毒拼接的速度由30-60分钟缩短到15-30分钟。全民抗疫，病毒基因测序和疫苗/药物研发是看不见的战场，在工信部的倡议中，这两项压轴，也给人工智能行业未来的攻关方向提出了新的建议。

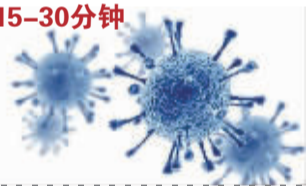


### ◇ 人工智能抗疫成绩单

顾斐与团队把病毒检测速度由**数小时**缩短到**半小时**



病毒拼接的速度由**30-60分钟**缩短到**15-30分钟**



浙江疾控、阿里达摩院、杰毅生物合作前基因检测时间**以天计**



浙江疾控、阿里达摩院、杰毅生物合作后基因检测时间**十几个小时**

新冠病毒基因检测一次可以检测**20个**样本平均每个样本**30-40分钟**



### ▶ “第一次现场支持，比‘996’辛苦” ◀

1月28日（大年初四），阿里达摩院主动联系了浙江省疾控中心。“我们和合作伙伴（杰毅生物）想要各捐一套设备过去，帮疾控中心搭建全基因组检测分析平台，目的是加快分析速度。”顾斐回忆。

其实自疫情暴发，这位哈佛大学医学院博士后与他的团队就保持着密切关注。他的主要工作是用人工智能的方法来研究疾病与基因突变的关系，目前在阿里达摩院医疗AI团队负责基因方向研发。这种特殊的交叉，让顾斐和团队一直希望能有机会去帮助

一线人员。

在对接后的第四天，浙江省疾控中心、阿里达摩院和杰毅生物上线了自动化的全基因组检测分析平台，把疑似病例基因分析的时间从数小时缩短至半小时。

据顾斐介绍：“我们在基因研究和算法上一直有技术储备，去现场之前已经准备好了基因测序方案”。一个小插曲是，因为涉及到数据保密等问题，阿里云等后方团队在正式合作前期，还做了搭建网络、服务器组建等工作，先后有上百人提供了支持。

这种短时间内多部门配合并不多见，在后来的实际合作中，顾斐和团队也遇到了很多第一次和意料之外。

“这是我个人第一次现场支持，之前大部分只需要在线上负责数据分析。”顾斐说。从1月28日开始，他每天要处理疾控中心以及公司的事情。工作时间从早上八点半到次日凌晨一两点，每天家和疾控中心两点一线。

谈到工作状态的变化，顾斐坦言，“现在的工作时间比‘996’要长，刚开始不适应，有点累”，但也觉得“小兴奋，因为很有责任感”。

### ▶ “之前基因检测以天计，现在十几个小时” ◀

这样的工作状态还要持续一段时间，因为人工智能与基因领域都有极强的专业性，这个工作非顾斐莫属。

具体到工作，系统调整是主要的内容，也是他和团队这次遇到的意外。

据了解，自动化全基因组检测分析平台是由阿里达摩院联合杰毅生物共同研发，可以给浙江省疾控中心在新型冠状病毒疫情防控上提供全自动建库和分布式计算分析能力。

其中杰毅生物开发了全自动高通量测序建库仪，把整体常规人工需要12个小时的工作时间缩短到2个小时。

然后是测序，阿里达摩院负责计算测序产生的海量数据，将病毒检测的整体速度由数小时缩短到半小时，把病毒拼接的速度由30-60分钟缩短到15-30分钟。

之前阿里达摩院和杰毅生物在各自领域都有一套自己的方案，现在三家一起来做，就要不断地改进。

据顾斐透露，现在三方的方案已经统一，但是还需要更多的数据支撑，系统仍在不断优化中。

顾斐向北京商报记者举例：“之前疾控中心有自己的全基因组测序方

案，有一套试验制配标准。现在三方要协调的包括选择哪个试剂、基因要双端测还是单端测、样本的浓度等。一开始这些都是不统一的”。

对比浙江省疾控中心原有的基因检测模式与新模式，顾斐说：“用之前的方式检测，有一些流程需要人工操作，比如把一台机器上的数据用人工拷贝的方式放到另一台机器上进行下一步，检测时间要以天计。现在利用新的平台，机器间的数据可以直接传输，加上算法优化，整个检测时间缩短到了十几个小时”。

### ▶ “希望重视更多基因研究子方向” ◀

“战疫”尚未结束，但已经给人工智能企业带来新的思考。北京商报记者注意到，这次行业代表性企业的支持，集中在AI测温、机器人咨询、无人配送等，少有像阿里达摩院这样的临床方向。

在杭州电子科技大学副教授徐伟栋看来，当代人工智能几乎已经是机器学习甚至深度学习的代名词。在机器学习领域，最大的成本与障碍不在算法，而在足量的训练样本。在基因领域，医学与生物学最大的成本与障碍不在计算上（基因测

序除外），而在将基因、蛋白组学的结果与临床结果的联立上，这需要大量的临床试验，很慢、很费钱，甚至涉及伦理困境。

具体到基因方向，中国人工智能学会会员张孝容认为，人工智能大致有3个阶段，基因预测、器官再造、基因编辑。目前企业的研究和探索主要在第一个阶段，突破本阶段可能需要很长的时间，需要投入的很多。一方面是因为AI基因领域的相关人才少；另一方面是回报难料，商业化较难。

谈及此次疫情对人工智能企业的意义时，顾斐直言：“阿里达摩院基因研究还会进一步扩大，接下来希望进一步招募更多的专业人才”。

顾斐已经想好了三四个子方向，比如蛋白结构预测、制药方面。原因是“这些方向更偏临床，会有实实在在的在帮助。现在很多机构都在根据新型冠状病毒的结构进行药物的筛选，其实阿里达摩院也在做这个工作”。

北京商报记者 魏蔚  
图片来源：企业供图