

# 中国科学院院士陈润生： 解析生物医学大数据，人工智能是最适用手段

## 院士访谈

◎本报记者 张佳星

近段时间，人工智能大模型展现了强大的自然语义处理能力，再次掀起了人工智能的研究热潮。从图灵测试开始，机器人能否与人对话就被用作判断机器是否具有高级别智能的依据。大模型的建立，让人工智能距离通过图灵测试更近一步。

那么，当前的人工智能还有哪些潜力？在生物医学领域应如何更进一步发展适用的人工智能？科技日报记者日前对中国科学院院士、中国科学院生物物理研究所研究员陈润生进行了采访。

### 生物信息数据量大、维度广

记者：距离人类遗传密码被破译已经过去了20多年，但似乎仍有很多健康问题难以解决，这是为什么？

陈润生：过去我们认为生物网络是“单色”的，或者说单因素作用的。因为当时我们认为具有生命功能的分子只有蛋白质，遗传物质经过转录、翻译的过程转化为蛋白质，进而执行生命活动中的各项功能。

基于这样的认知，人类遗传密码的破译被理解为解读“生命天书”。将人类全部30亿个碱基序列全部测序完成，就能读懂这部“生命天书”，提纲挈领地搞清生命活动的运转体系，从而揭开所有健康的秘密，获得解决健康问题的手段。

但现在我们发现，生物网络不是“单色”的，生命活动的调控、抑制、互作中，还有核酸的参与。比如有些核酸序列也会发挥之前被认为属于蛋白质的功能，调控细胞活动。在承载人类遗传信息的30亿个碱基对中，还有很大一部分虽然我们读出了序列，但并不知道功能，它们在生命活动中究竟起到什么作用也没有被研究清楚。大量这样的区域被称为遗传物质的“沙漠区”……因此，当前对人类遗传密码的破译进程只了解决健康问题过程中的一小步。

记者：既然生命活动这么复杂，是否需要更多的生物信息数据才有希望破解健康难题？目前需要充分利用的有哪些维度的生物信息学数据？

陈润生：这30年来，人类遗传密码的破译带动着整个生物医学领域内广泛的数据都成为了大数据。

生物信息学最早的研究对象是遗传密码，研究任务是对遗传密码进行收集、整理、存储、发布、分析和解释。现在数据更广泛了，生物信息学的研究从20世纪八九十年代创立时的无信息可用，到现在数据呈现海量、多维、异质化特征。比如通过电子病历大数据挖掘，我们可以得到很多重要信息。中山大学一些研究人员通过电子病历分析找到的科学结论被发表在了国际著名期刊上。又比如可穿戴设备带来的生理生化指标，这类信息具备实时、环境情绪相关等特性。这些数据还包括医院中的影像学数据、临床病人在治疗用药后的病情变化数据、环境数据、微生物数据、地质数据、化学农药残留数据等。所有这些跟健康相关的数据集纳起来，构成了非常复杂的健康大数据类型。

### 可大幅提高生物信息处理能力

记者：如何解析复杂的健康大数据？

陈润生：健康大数据很复杂，举例来说，电子病历写的字，处理这样的信息需要自然语言处理模块，生理生化指标如脑电、心电等信息需要处理的是波形图，而影像学要处理图像，组学要处理符号。对这些性质各异生物信



图为陈润生院士。 陈超/中新社/视觉中国

息进行全面提取处理，就如同秦始皇统一六国不仅要统一文字、语言，还要统一货币、度量衡，需要将不同的信息融通，再做解析和决策。

当前最核心的生物医学大数据问题是解决各种可测量的数据，包括遗传密码与表现型之间的关系问题。我们知道分子水平的基因组学、蛋白质组学研究产生的数据量不计其数，近年来发展起来的表型组学研究也在不断积累数据量，我们会逐渐拥有大量的表现数据。那么这两大数据如何偶联？如此规模巨大、系统复杂的分析只能由人工智能来完成。

目前来看，要解析生物医学大数据，人工智能技术是最适用的手段。

记者：从最初的人工智能到阿尔法狗再到大模型，人工智能发生了哪些变化？当前的大模型在处理生物医学信息方面有哪些优势？

陈润生：20世纪80年代，由于受算力和算法的限制，我们能设计运行的人工智能神经网络非常简单，只有几层。

现在的人工智能神经网络，比如阿尔法狗、阿尔法折叠等，它的网络模型架构达到200—300层。当今的计算能力能够使它完成更加复杂的任务。

人工智能大模型让人工智能更进一步。我认为大模型的出现表明人工智能有了多方面、更深刻的进展。第一方面是对自然语言的处理能力发生了根本性的变化。处理人类的自然语言对于计算机来说很困难，突破这个能力后，人工智能可以拥有全局性的知识作为计算的基础。

第二方面进展是具备了预学习能力。人类之所以能产生智慧，是因为从出生以来一直在不断积累学习。现在在大模型中，计算机已经可以预学习了，能够在大量的知识被预先获得的前提下，再将具体科学问题交给计算机，而不是针对某一个特定问题进行学习。

第三方面进展是可积极促进模态融合。模态融合可以这样理解，比如人工智能不仅能完成基础医学里的结构预测工作，还能完成临床医学中的识图判读工作，更重要的是可以把这两个能力融合起来，即人工智能在每个领域都具备专家的知识，又能够融会贯通，举一反三。

基于以上三点，我认为发展人工智能作为未来医学大数据挖掘的工具值得重视。

### 以人工智能解析生物医学大数据

记者：很多人在健康出现问题之后的第一反应是“为什么会这样”，以后人工智能可以解答这样的问题吗？



当前最核心的生物医学大数据问题是解决各种可测量的数据，包括遗传密码与表现型之间的关系问题。我们知道分子水平的基因组学、蛋白质组学研究产生的数据量不计其数，近年来发展起来的表型组学研究也在不断积累数据量，我们会逐渐拥有大量的表现数据。那么这两大数据如何偶联？如此规模巨大、系统复杂的分析只能由人工智能来完成。

陈润生：想要回答“为什么会这样”这个问题涉及到人工智能界的两个提法：弱人工智能和强人工智能。弱人工智能就像阿尔法狗，能处理特定任务。而强人工智能被认为是有可能制造出来的、真正能推理和解决问题的智能机器。

所以你的问题其实是在问，强人工智能有可能出现吗？过去我认为是不可能的，但现在看来，我觉得随着大模型的出现，一些问题正在破冰，强人工智能有了出现的苗头和希望。当然，现阶段人工智能的发展与真正的强人工智能还差得很远。

当前我们正在进行一些与医学相关的人工智能研究。比如扩大健康问题相关的参数，抽提面部、表情等特征，也融合基因组、转录组、蛋白组、宏基因组、表观组等全部维度的特征，甚至加入了中医内容。通过多模态融合建造医学多模态数据智能整合平台，我们称之为“灵枢”，并将其用于医学健康问题的实践。

记者：网上有人调侃，能写论文的人工智能却无法解释清楚“驴肉火烧”“鱼香肉丝”，而生物医学领域对于人工智能的要求更高，您认为如何让人工智能具有更严谨的运行能力？

陈润生：如何才能让人工智能解释清楚“驴肉火烧”“鱼香肉丝”等人类世界的问题，其实是在问如何才能让人工智能达到强人工智能的水平。

人工智能的智能水平主要由三个因素决定：数据、模型和算力。数据掌握在用户手中，模型需要不断研发讨论精进，算力是由计算机的能力决定的。

从这三个决定因素入手，我们能够发现，在数据方面，要解决相应的科学问题，人工智能掌握信息量高、信噪比足够的信息就可以了。在算力方面，要依靠计算机科学家不断从硬件等解决方案入手提高运算能力。

只有模型方面是值得研讨的，现在的模型理论不够达到强人工智能的要求，如何让模型更强？

目前有几个途径可以让模型更强，即增加知识量、增加网络节点数、增加网络复杂度。当前国际上几个大模型的处理数据量已达到万亿参数级别，在此基础上，我认为需要进一步增加的是网络复杂度。

我们可以从生理学家的脑生理切片模型中得到启发。新生儿的神经网络很简单，复杂度不够，学不了多少东西。成长让人的神经网络逐步复杂，进而承载更多内容。因此，我们可以从研究脑发育中获得启发，进而形成一个更完善的理论基础框架，获得适用的模型，使人工智能走向强人工智能。

任、网络服务提供商等多种责任。因此，有关部门进行短视频直播带货的监管要更为复杂。”

### 分级分类管理利于净化环境

《报告》指出，2022年我国的直播带货消费维权舆情主要反映了产品质量、虚假宣传、不文明带货、价格误导等8方面的问题。其中，产品质量和虚假宣传仍然是直播带货面临的两大主要问题。

陶盈指出，对于直播带货过程中出售的产品性能不符合国家或者行业标准的情况，消费者可以请求产品生产者与销售者承担侵权责任；如果直播过程中存在虚假宣传的问题，消费者可以要求广告经营者、发布者或有关带货人员承担连带责任。

推动直播带货的高质量发展，加强监管无疑是重要手段之一。陶盈认为，由于直播带货的门槛低、规模庞大且违法行为隐蔽性强，监管治理难度较大。

《报告》建议对实施带货行为的直播间实行分级分类管理。在陶盈看来，这样的分级分类管理对于加强监管和治理很有帮助。“目前直播带货领域良莠不齐，平台规模、直播类型、产品或服务内容、营利方式、

付费模式、用户对象、粉丝数量、社会影响力等因素差异性极大。”陶盈表示，“对直播间进行分级分类管理，确定重点监管领域和监管对象，有利于净化网络环境。而对涉及违法违规的直播账号，根据其影响或危害程度进行暂停直播或封禁等处罚，则有利于提高监管效率，有针对性地开展管理。”

陶盈指出，监管部门应当建立常态化的监管机制，通过定期抽查直播内容，对直播带货产品抽样抽检。还可以设置有奖举报机制等，科学治理直播带货乱象。有关部门也可以积极研发技术监管手段，通过视频、语音、图片自动识别技术，及时、主动发现直播带货中的违法、违规行为。

广西社会科学院社会学研究所所长姚华此前同样表示，监管部门、平台公司以及社会各界应该共同努力，促进直播带货行业的健康有序发展，持续为电商经济高质量发展提供强劲动力，为消费者带来更好的购物体验。

在陶盈看来，规范直播带货新业态发展，需要在保障新经济健康发展、经营者营业自由的同时，加大对于违法乱象的监管惩戒，形成网信、市场监管、广电、公安等部门之间的信息共享、协同联动，合力让直播带货行业真正行稳致远。

## 热点追踪

### 我国163个矿种中 近四成储量上升

科技日报讯（记者操秀英）科技日报记者6月14日从自然资源部例行发布会获悉，根据2022年度全国矿产资源储量统计数据，已有查明矿产资源储量的163个矿种中，近四成储量均有上升。锂、钴、镍等战略性新兴产业矿产储量分别同比增加57%、14.5%和3%。

自然资源部矿产资源保护监督司司长薄志平介绍，锂被誉为绿色能源金属和白色石油，广泛应用于储能、化工、医药、冶金、电子工业等领域。随着全球绿色低碳转型和新能源汽车快速发展，锂资源的战略地位日益凸显，各国纷纷将锂列入其关键矿产目录，高度关注锂资源供应安全。

薄志平说，我国锂矿种类丰富，有盐湖卤水锂矿、锂辉石矿和锂云母矿，但总体品位较低，优质锂资源较少，主要分布在江西、青海、四川和西藏等4省（区），河南和新疆也有少量分布。

2022年度统计数据表明，江西锂矿储量超过青海和四川，居全国第一，占全国总量的40%。2022年度全国锂矿储量增量也主要在江西，占增量的94.5%，来自江西宜春市的两个锂矿山。

薄志平表示，宜春锂矿床类型主要为锂云母型，多为低品位含锂的陶瓷土矿。由于覆盖层薄，大多采用露天方式开采。随着锂云母提锂技术提升，成本大幅下降，同时锂电新能源产业兴起对锂矿的需求也上涨，市场开发意愿大幅增强，宜春的锂矿在国内快速释放锂资源供给增量方面优势明显。但是随着含锂的陶瓷土矿开发利用规模快速增长，长石粉和锂渣的产出量将大量增长，科学处置、综合利用长石粉和锂渣已成为当地锂电新能源产业发展需要特别关注和解决的问题。

会议还介绍了我国其他矿产资源的储量情况。2022年度统计数据显示，在油气矿产方面，2022年剩余技术可采储量均有不同幅度增长，其中石油同比增长3.2%；天然气同比增长3.6%；页岩气同比增长3.0%；煤层气与去年基本持平。2022年度有3个油气田新增石油储量规模达到大型；有6个油气田新增天然气储量规模达到大型。

非油气矿产方面，铜、铅、锌等大宗矿产储量分别同比增加16.7%、7.1%和4.2%；铍、锆、镓等稀有金属储量分别同比增加11.7%、7.9%和16.5%；金矿同比增加5.5%；非金属中，普通萤石、晶质石墨同比增加27.8%和3.5%。

薄志平介绍道，矿产资源储量统计是自然资源统计工作的重要组成部分，是掌握资源家底、支撑自然资源高质量发展、保障能源资源安全的重要基础性工作。2020年5月1日，我国实施了新的矿产资源储量分类国家标准。3年来，自然资源部作出了一系列工作部署，构建了新分类标准下的统计制度，完成了矿产资源储量新老分类数据转换，对储量库进行了全面清理，实现了两个全覆盖。通过建立常态化的质量监控机制，按年度对统计数据开展全面核查、重点检查和综合评估，建立问题台账，“抓源头、补短板、遏增量”，数据及时性、全面性和准确性得到了显著提升，新分类标准下的资源储量数据框架和体系已全面形成。

## 8部门发文推进深远海养殖

新华社讯（记者于文静）农业农村部、工业和信息化部、国家发展改革委、科技部、自然资源部、生态环境部、交通运输部、中国海警局近日联合印发关于加快推进深远海养殖发展的意见，提出全产业链全环节加快推进深远海养殖发展的重点任务。

这是记者6月12日从农业农村部了解到的消息。据了解，意见坚持市场主导、科学布局、科技引领、绿色生态、安全发展原则，围绕深远海养殖发展的重点领域和关键环节，提出加快推进深远海养殖发展的重点任务，是我国首个关于深远海养殖发展的指导性意见。

意见提出，要优化养殖空间布局，发展深远海养殖要符合国土空间规划和养殖水域滩涂规划要求，桁架类网箱及养殖平台、养殖工船原则上布设在低潮位水深不小于20米或离岸10公里以上的海域，重力式网箱布设海域水深不小于15米。要推动全产业链发展，大力推行健康养殖，做精做好水产种业，积极发展水产品加工流通，培育区域性优质品牌，鼓励创建深远海养殖类型的国家级水产健康养殖和生态养殖示范区。

意见指出，要加强生态环境保护，深远海养殖项目要依法开展环境影响评价，严格落实生态环境保护相关措施要求。要推进产业技术创新，加大对先进养殖技术和设施装备研发支持，完善技术体系和规范标准，提高深远海养殖的安全性和经济性。要落实检验登记要求，养殖工船依照《中华人民共和国渔业船舶登记办法》，按照养殖渔船进行登记，重力式深水网箱、桁架类网箱及养殖平台由地方人民政府指定的主管部门制定规则进行登记。

意见提出，要强化安全生产措施，深远海养殖生产经营者要落实主体责任，遵守相关规定，做好风险防范。要健全监管协作机制，建立由同级农业农村（渔业）部门牵头，各有关部门共同参与的深远海养殖监督管理工作协作机制，明确职能分工，形成齐抓共管的监管格局。



图为在福建省连江县定海湾内的“振渔1号”深远海养殖平台上，工作人员展示养殖大黄鱼。 新华社记者 姜克红摄

# 开展针对性监管，让直播带货走上正轨

◎实习记者 李诏宇

“点一下就能参与限量秒杀活动，家人们快来下载……”

对于经常使用电子商务平台的人来说，上述场景一定不陌生。近年来，随着电商活动的日益火爆，直播带货正成为广大消费者购物的一个重要选择。然而，直播带货背后也暗藏风险。

消费者网、对外经济贸易大学消费者保护法研究中心、北京阳光消费大数据研究院等有关机构前不久联合发布了《直播消费维权舆情分析报告》（以下简称《报告》），对直播带货中存在的各种问题做出了分析。

那么，公众关切的直播带货中存在的问题应如何监管？记者采访了相关专家。

### 短视频直播带货监管更复杂

直播带货指的是网络主播在进行直播娱乐活动的同时进行有关商品的推介或售卖的一种行为，往往可以分为传统电商平台的直播带货和短视频直播带货。

尽管直播娱乐活动在我国兴起较早，

但那时的直播娱乐活动主要依赖打赏等手段营收，经常出现所谓“有流量，没收入”等尴尬情况。某种程度上说，直播带货正是为了解决上述问题而生。

在首都经济贸易大学法学院副院长、副教授陶盈看来，直播带货不仅代表着新消费形态的崛起，也是挖掘经济新蓝海、激发消费活力和内需潜能的重要手段，具有不容否认的积极意义。

不同的平台也为直播带货带来了不同的特点。“相对于传统电商平台直播带货，短视频直播带货在带货模式、监管方法等方面均存在差异。”陶盈表示。

具体来说，短视频直播带货在模式上要更加灵活，不仅商品厂家或者大型经销商会带货给消费者，专职的主播、网红等也会参与带货。此外，为吸引消费者，短视频直播带货往往会采取情景式演出等方式进行，而非简单的广告式推介。

“短视频直播带货灵活的模式有助于增强消费者的黏性和信任度，进而促进商品的营销。但其特点也导致对它的监管更加复杂。”陶盈指出，“短视频直播带货涉及主播、直播账号运营方、品牌方、经销商、生产商、直播平台等多方主体，可能涉及产品质量、虚假宣传的广告经营者和代言人责