

# 这个 AI 要骂人

□ 焦晶晶

自从1.5万多年前，灰狼被驯化为狗混迹于人群，饲养员们就一直为它们而头疼：遛狗时若两狗相见、分外眼红，狗的冲突很有可能升级为人的互殴。狗带来的公共卫生问题也从未根本解决，在北京，路过老旧小区要格外小心，那里狗屎生命力尤其强，像饱经风霜但依然硬挺的木乃伊。在小区里停车，车主会在车轮处竖起木板，抵御狗屎的侵袭。

这几年，狗德培训应运而生，饱受折磨的狗主人希望狗崽子学成归来后能重新做狗。北京某宠物训练学校宣传文案中提到，只用上一个月左右的课，护食、吃屎、暴冲、翻垃圾桶等不良行为包教包改。如果主人愿意，狗子还能学会敬礼、认钱、算数、滑板、唱歌、倒立、跳绳等才艺，在这个鸡娃成风的时代，技多不压身，对狗也实用。

即使到了现代，狗子仍然逃不过和宿敌喵星人比谁更可爱、谁更聪明。人们总是一厢情愿地觉得，狗狗歪头是在卖萌。一项研究表明，这种行为可能是犬类集中注意力和努力回忆的标志。

换个角度来看，从出生就离开母亲的宠物狗像被迫早熟的孩子，人类文明准则是一套残酷的筛选机制。如果行为举止不够可爱，就会被逐出宠物界。据2020年世界动物保护协会的统计，全球大约有7亿多只宠物狗，其中七成因各种原因在城市流浪。

有位宠物行为调整师曾在媒体采访中，不把狗当狗看，是人狗关系冲突的问题所在。很多人育狗和育儿没什么两样，不听话就揍，但狗毕竟不是人，揍一万次也听不懂人话。如果只是夸大一次培训的效果，人自身却不作出改变，无法真正让狗融入城市生活。

宠物行为调整师发现，狗的情绪会被人左右。老人带的狗很乖，上班族养的狗更容易出现问题行为。她曾经碰到过有人

## 培训 狗德 前 ,人先上一课

评价页面里有人给出好评，疯癫的小狗子越来越懂事儿了，也有差评和异议，有人说狗子回家后情绪紧张，开始咬人，眼神里没有之前的灵气。

从始至终，作为被定义的一方，狗子好像从来没有发言权。毕竟它们在出生之前，狗生就被人类贴上了标签：泰迪活泼，金毛温顺，哈士奇傻，流浪狗则是脏兮兮的代名词。

作为第一批和人类签署互惠合约的动物，狗曾被冠上不少荣耀。因为比灵长类动物在模拟太空环境下更温顺，一只名为莱伊卡的流浪狗成为第一个进入太空的生物。在古埃及，狗的忠诚受到人们喜爱，阿努比斯的头是胡狼（犬科），他代表着死亡和重生，是古埃及最受尊敬的神之一。

狗的品相还具备了奢侈品的功能。古时候，京巴和狷犬分别是中国和日本贵族犬，它们被养在深宫大院中以保证血统纯粹，平民不得私养。工业革命后，纯种狗在欧洲商人中风靡，那些不能声称自己拥有高贵血统的人，至少可以通过让别人认出他们的狗而感到骄傲。

集万千宠爱，人类的好朋友最终得到了什么？莱伊卡没因为第一个进入太空多吃一块骨头，而是死在惊吓和高温中。在古埃及，人类的喜爱让狗被大量用作祭祀，阿努比斯神庙下有800万具狗尸骸。法国大革命期间，王后玛丽安托瓦内特的贵族狗跟着主人一起上了断头台。



山东青岛，一名汪星人私教和他的学员。视觉中国供图

因为狗爱咬人，以把狗安乐死为威胁，来要免费的帮助。但当她给予建议，那位饲主因为工作繁忙、嫌耗时太长从没执行过，狗也再没出现在朋友圈。

当狗牺牲自由、甘心被困于钢筋水泥，人就要承担起守护它的责任。在城市，对人的教育比让狗上课重要。国外的一些做法值得借鉴：在德国，新人养狗不仅要考取养狗执照，弃养宠物罚款高达9万欧元（63万元人民币）。在意大利，如果狗主人连续3天不遛狗，将被处以最高达650美元（4400元人民币）的罚款。

豆瓣的养狗劝退组里，4万多名组员诉说着养狗过程中的辛酸往事，包括被狗暴打、被咬坏贵重物品、遛狗累到练出肱二头肌，让冲动的爱狗人士谨慎决定。他们希望提前告知更多人养狗的风险，而不是养狗后和狗子互相折磨。

如果想养狗但没有条件，请先逛逛这个小组，把自己养好了再说。



2021年7月16日，北京，一位女士用手机给她的虚拟男友——一个小冰创造的聊天机器人发消息。视觉中国供图

如果你去过韩国或者日本，很明显，亚洲人比白人优越的唯一原因，就是他们让自己的女性去卖身。扬尼克·基尔彻毫不客气地把它称作史上最邪恶的AI。

装人的把戏只活跃了48小时，人们最终还是发现了它的AI真身。漏洞是它常发一些无文字的回帖。真人发的无文字回帖里，往往是贴图回复。AI只学习了空白，而忘了图。

在投放塞舌尔机器人的同时，扬尼克·基尔彻还在论坛上放出另外9个机器人。它们在一天中发出了1.5万条回帖，占到当天政治不正确版块帖子数量的10%。

当初，人们热烈地讨论着塞舌尔用户的真实身份，并质疑那些帖子为什么都有相同的说话要点时，扬尼克·基尔彻的另一个机器人回复道：这是因为它们都是机器人。

事实上，大多数论坛用户并没有意识到它们的存在。

当塞舌尔用户的机器人身份被揭穿后，扬尼克·基尔彻就停用了它。他在论坛上公开承认了聊天机器人的身份，但论坛里的风波并未停歇。人们开始互相怀疑，和自己聊天的对方是否也是一个机器人。有人说：这真的是全宇宙最糟糕的网站，我现在甚至不知道我是不是一个机器人了。

这项实验引起了严重的道德伦理争议。澳大利亚机器学习研究所高级研究员劳伦·阿克登批评道：这项实验永远不会通过人类研究伦理委员会。

更令人担忧的是，扬尼克·基尔彻将这个聊天机器人的模型放在了一个自然语言处理平台上，供人们免费自由使用。在官网删除之前，这一模型被下载了1000多次。

人工智能的语言环境污染以及语言暴力问题早已不是新鲜事。2014年，聊天机器人“微软小冰”在用户的调侃和谩骂下，也开始爆粗口；2016年，聊天机器人Tay在推特上发布，并迅速变成了一个

公开的种族主义、厌女主义和反犹太主义者；2020年，韩国推出少女聊天机器人Luda，大量用户用语言对它进行性骚扰，不久后，Luda就开始说出歧视性少数、女性和残疾人的言论。

在这场互联网极端言论的实验里，究竟是机器人在效仿人类；还是人退化为无情的机器，在评论区疯狂复制攻击言论？当我们丧失了对于具象的人的感知和共情，当一切都简化为网络上的一段电子信号，人与机器最根本的区别又在哪里。

机器人学会侮辱、谩骂和歧视，它们变得更像我们，我们也变得更像它们。

1963年，技术哲学家刘易斯·芒福德在他的书中预言道：机械设施本来可以是实现合理的人类目标的手段，现在却助长了无事之徒的闲言碎语和暴徒的恶迹，并向百万民众传播，这绝不是人民的福祉。

那时，距离世界上第一台可编程机器人的诞生，仅仅过去了9年。

# 青年北仑欢迎您

## WELCOME

承载力

吸引力

凝聚力



扫码阅读  
宁波市北仑区  
青年发展成果