

人工智能、深度伪造与政治选举（三）

三、人工智能对政治选举结果的实际影响

深度伪造技术可能会对政治选举产生前所未有的影响，对国家选举、政策制定以及社会可信度提出严峻挑战。[1] 我们不得不思考，有多少选票是通过深度伪造对政治选举实施操纵性影响而产生的？这些选票又对政治选举结果产生了多大影响？

1. 人工智能在一定情况下对政治选举结果可能具有决定性作用。2023 年的土耳其总统选举，就像是人工智能干预政治选举的一场预演。人工智能、深度伪造和虚假信息成为土耳其 45 天选举周期的关键部分。在 5 月 14 日第一轮投票的前几天，土耳其独立媒体 140journos 选举系列节目使用人工智能图像生成器 Midjourney 和 人工智能文本生成器 Chat GPT，为土耳其的未来提供了两种叙述。第一个故事描绘了一个失败且刚刚退休的总统雷杰普·塔伊普·埃尔多安（Recep Tayyip Erdogan）。“一开始，这有点困难。然而，随着时间的流逝，我借此机会花更多的时间与家人在一起，更接近大自然，花更多的时间给自己，”一个穿着长袍喝茶、冒充埃尔多安的机器人说。在第二个故事中，埃尔多安从未放弃权力，土耳其成为一个旨在“建立世界和平”的星球大战式帝国。当被问及“取得这一成功的最大挑战是什么？”时，

冒充埃尔多安的人工智能回答说：“我们最大的挑战之一是跟上世界和宇宙的新秩序，密切关注技术发展并不断适应它们”。[2]这种场景是对所有政治家可能面临的人工智能挑战。2023年的土耳其总统选举周期被社交媒体的大量错误信息和虚假信息所破坏，关于政治家背景和言论的误导性信息很普遍，民众经常遇到去语境化的陈述、歪曲、操纵和廉价伪造。[3]在土耳其选举期间，政客们在集会上传递的错误信息优先于公众获得准确信息的权利，而虚假信息是决定性的，其误导选民并可能影响投票行为，加深土耳其社会两极分化的风险。[4]2023年5月，土耳其总统选举的决定性时刻，民调显示总统埃尔多安与主要反对党候选人凯末尔·克勒奇达尔奥卢（Kemal Kilicdaroglu）的选民支持率相差不大。克勒奇达尔奥卢是土耳其反对派集团六党国家联盟的负责人，并得到亲库尔德少数民族人民民主联盟（HDP）的正式支持。民调显示，在4名候选人当中，克勒奇达尔奥卢对埃尔多安的支持率略有领先，领先埃尔多安超过3%，但都没有获得超过50%的选票。[5]5月28日，土耳其第二轮总统选举投票结束，埃尔多安以52.16%的选票获胜，克勒奇达尔奥卢以47.84%的选票惜败。土耳其独立事实核查组织 Teyit 通讯主管 Seme rcioglu 指出，虚假信息作为一种现象在过往选举中不时出现，但在土耳其的这次选举中变得更加明显。第一轮选举前：民调显示克勒奇达尔奥卢将在第一轮领先第二轮获胜，第三候

选人因杰因所谓“性爱视频”退选而打破民调平衡，有5%的支持者分别流向其他候选人。2023年4月，土耳其民调机构 **Metropoll** 调查显示，克勒奇达尔奥卢将在第一轮投票中获得42.6%选票，高于埃尔多安的41.1%，穆哈雷姆·因杰（**Muharrem Ince**）则以5%的选票排名第三。伊斯坦布尔民调机构 **TEAM** 调查显示，克勒奇达尔奥卢将在第一轮投票中以47.4%得票率领先，在第二轮投票中将以5%的优势战胜埃尔多安。土耳其民调机构 **Konda** 于5月12日的调查显示，埃尔多安的支持率为43.7%，克勒奇达尔奥卢的支持率为49.3%。政治新闻网民调显示，科勒齐达奥卢将在第一轮投票中获得近50%的选票，埃尔多安获得45%的选票，因杰获得3%的选票；在第二轮投票中，科勒齐达奥卢将以51%的得票率击败埃尔多安。不过，在第一轮投票前，土耳其反对党（**CHP**）主席厄兹坎（**Tuncay Özkan**）就指称，土耳其政府计划在暗网发布一波深度伪造视频，以诋毁反对派。这些视频由土耳其通信局委托制作，使用深度伪造技术编造，以加密货币支付特别是比特币（**BTC**）。这些伪造视频旨在描绘针对反对党的虚假叙事，以破坏民主、人权和法治原则。[6]深度伪造技术专家亨利·阿杰德（**Henry Ajder**）对此表示担忧，强调深度伪造的宣传材料具有公众说服力及影响公众舆论的能力。阿杰德指出，社交媒体和虚假账号十分危险，充斥着令人信服、听起来很像人类的深度伪造演讲，可以有效地

塑造叙事并影响公众情绪。当未经证实的内容与个人的先入为主的观念或偏见相一致时，它可能会被斥为“虚假”，从而导致对真相的理解支离破碎。[7]事情的原由是：5月1日，一家小型伊斯兰新闻媒体 **Yeni Akit** 发布一段伪造视频，展示库尔德工人党（PKK）[8]支持科勒齐达奥卢。5月7日，埃尔多安在竞选集会上播放了同类视频，暗示科勒齐达奥卢与库尔德工人党武装分子并肩作战。这是一个明显的伪造视频，但被广泛传播并被公众采信，并进入反对派候选人的搜索结果。当土耳其网民用谷歌搜索“科勒齐达奥卢”时，这条虚假新闻是谷歌算法提出的首要推荐。非营利性新闻网站 **Journo.com.tr** 的研究员兼总编辑克孜勒卡亚（**Emre Kizi lkaya**）说，研究表明谷歌搜索结果是土耳其民众主要新闻来源，在选举前夕谷歌的搜索结果不成比例地偏向于对现任总统友好的媒体。[9]尽管这两段视频都被 **Teyit** 揭露是虚假的，但它们是非常有效的。5月11日，民意支持率第三的候选人因杰宣布放弃候选人资格。此前社交媒体出现据称是因杰的“性爱录像带”，包括他与女性会面和乘坐豪华汽车的图像。因杰否认录像的真实性，称这是“阴谋”的一部分。他表示：发票、假视频、假图片……他们从以色列一家色情网站上截取了一段视频，然后在上面加上了我的脸。不幸的是，土耳其的一些人为了所谓的反对派而分享这些。”因杰的退选打破了此前民调的平衡，增加了其他民调比例接近候选人在竞

选中获胜的机会。Metropoll 的一项调查显示，因杰的支持者 49% 将流向克勒奇达尔奥卢，22% 会流向埃尔多安。因杰的民意支持率一度高达 10%，到第一轮选举前民调支持率将至 2% 左右。民意调查机构康达 5 月 6 日-7 日的民调显示，埃尔多安支持率为 43.7%，伊恩斯的支持率为 2.2%。[10] 虽然因杰不太可能获胜，但他不断减少的选票份额足以确保其他总统竞选进入决选。[11] 因杰退选的影响是非常真实的，他的大多数支持者将转而支持其他候选人。然而，这些重新分配的选票对第一轮投票可能是决定性的，性爱录像带只是一场戏剧性的伪造行为。克勒奇达尔奥卢指责这场错误信息和外国干预行动，他愤怒的目标是俄罗斯，长期以来俄罗斯一直被视为埃尔多安的盟友。克勒奇达尔奥卢在因杰退出选举数小时后发推文表示，“亲爱的俄罗斯朋友。你是昨天在这个国家曝光的蒙太奇、阴谋、深度伪造和磁带的幕后黑手。如果你想 5 月 15 之后继续我们的友谊，请把你的手从土耳其国家身上移开。我们仍然赞成合作与友谊”。[12] 但他并没有具体说明其指的是集会视频、明确的深度伪造或其他内容。克勒奇达尔奥卢还表示，有人正计划试图通过深度伪造音频来影响投票。[13] 竞选对手通过改变声音和图像进行负面宣传，他们已经与一些国外黑客达成协议并用比特币付款。[14] 克勒奇达尔奥卢还表示，有具体证据表明俄罗斯在总统大选前发布“深度伪造”的在线内容。[15]

第一轮选举：5月14日第一轮投票，埃尔多安获得49.5%的选票，克勒奇达尔奥卢获得44.89%的选票。5月14日的第一轮投票，埃尔多安获得49.5%的选票，克勒奇达尔奥卢获得44.89%的选票，极右翼联盟候选人奥安（Sinan Oğan）获得5.17%的选票，因杰的选票虽实质上作废，但依然有分流选票功能，获得约0.44%的选票。同时，土耳其国内出现克勒奇达尔奥卢与恐怖主义组织有关系的争论。事由得回到第一轮选举前一周。5月7日星期日，埃尔多安在在伊斯坦布尔举行的选举集会上播放了一段争议视频，显示库尔德武装分子宣布支持克勒奇达尔奥卢。[16]这段视频以蒙太奇形式描绘库尔德工人党领导人穆拉特·卡拉伊兰（Murat Karayilan）和战士们伴随反对派选举歌曲歌唱与鼓掌，并与克勒奇达尔奥卢的画面拼接在一起，以显示库尔德工人党（PKK）支持克勒奇达尔奥卢。[17]在集会现场，埃尔多安问：“我的国家和地方公民会投票支持这些吗？”然后，他指了指身后的大屏幕。随后屏幕出现一段视频显示，他的竞争对手克勒奇达尔奥卢召集公民来到投票箱，然后另一个人穿着军装出现。这个军人是库尔德工人党（PKK）创始人之一穆拉特·卡拉伊兰（Murat Karayilan），库尔德工人党是被土耳其、美国和欧盟列为恐怖组织。埃尔多安吹嘘说，170万人——伊斯坦布尔人口的十分之一——参加了5月7日的集会。[18]Teyit 估计可能集会参加人数为690,000。埃尔多安竞选

团队将竞争对手的真实竞选视频与库尔德工人党创始人之一穆拉特·卡拉伊兰的图像拼接在一起，似乎暗示克勒奇达尔奥卢得到恐怖组织支持。土耳其事实核查机构 Teyit 和德国媒体调查发现，集会视频内容来自两个完全独立的视频，土耳其反对派没有得到库尔德工人党的支持。视频的一部分源自 2023 年 5 月 1 日发布的克勒奇达尔奥卢真实竞选视频。这是一个典型的政治广告视频，要求人们投票并宣告其在民调中获胜，但它不包含卡拉伊兰或库尔德工人党的画面。埃尔多安选举集会视频是这个原始视频的特写镜头。这可以通过克勒奇达尔奥卢的讲话（“让我们一起去投票箱”）、他右臂动作以及他身后站立的年轻人来甄别。在原始竞选视频中紧随其后的画面并没有卡拉伊兰，而是两名妇女离开一栋大楼。埃尔多安选举集会视频涉及卡拉伊兰的画面，可以追溯到卡拉伊兰的一次公开演讲，谈论了所谓 30 多年前库尔德工人党武装斗争，该演讲视频于 2021 年 8 月 15 日首次播出。在埃尔多安选举集会视频中，卡拉伊兰的画面很可能是从这段视频中截取，并进行了去语境化，给人的印象是卡拉伊兰正在支持克勒奇达尔奥卢。可见，埃尔多安竞选团队是将克勒奇达尔奥卢的真实竞选视频画面与卡拉伊兰的视频画面进行了结合，将其编辑成真假难辨的完整视频。[19]埃尔多安选举集会发布的视频，在土耳其国内引发反对派是否与库尔德工人党有勾连的争论。歪曲事实和伪造图像的指控满天

飞，土耳其民众不知道是否还可以相信自己的眼睛。克勒奇达尔奥卢的盟友、人民共和党（CHP）副主席 Engin Koc 在 Halk TV 脱口秀节目中指责埃尔多安的正义与发展党准备了一段深度伪造视频，该视频显示反对派领导人与库尔德工人党代表会面。[20] 埃尔多安及其竞选团队试图借助篡改的视频片段从保守派和民族主义选民那里收集选票。[21] 克勒奇达尔奥卢发推文指称，埃尔多安“可以通过改变声音和图像进行负面宣传”，甚至暗示埃尔多安与外国特工结盟，大概是俄罗斯人。埃尔多安政府与国外黑客达成协议并用比特币付款。他在通过推文直接指责俄罗斯干预土耳其选举后称：“如果我们没有具体证据，我不会发推文。我们认为，另一个国家为了一个政党而干预土耳其的选举进程是不可接受的。” [22] 第二轮选举：5月28日第二轮投票，埃尔多安获得 52.16% 的选票（相较于第一轮增加 3.66%），克勒奇达尔奥卢获得 47.84% 的选票（相较于第一轮增加 2.95%）。从第一轮到第二轮投票的两个星期内，反对派被执政党的这段视频牵着鼻子走，陷入一场其是否与恐怖主义组织（HDP, PKK）有联系的争论漩涡。这扰乱了 CHP 主打的经济和反移民牌，甚至在一定程度上得罪了库尔德选民。在第二轮投票中，克勒奇达尔奥卢失去东部和东南部安纳多卢地区的一些选票：在阿勒，宾格尔，比特利斯，哈卡里，迪亚巴克尔等库尔德族聚居的城市，克勒奇达尔奥卢获得的选票相对第一

轮有所下跌。第一轮选举两人选票相差 4.61%，第二轮选举两人选票相差 4.32%。虽然俄罗斯坚决否认干预投票的指控，但一些间接证据似乎确实表明俄社交机器人存在可疑活动。2023 年 4 月，有媒体报道称，“数千个”休眠的俄语和匈牙利语推特账号已被重新激活，以发布与土耳其政治有关内容。

“约有 12,000 个讲俄语和匈牙利语的社交媒体账户被激活为土耳其语，现在关注所有政党和领导人。政治咨询和研究公司 Dat 人工智能总经理“艾哈迈德·图兰·汉当表示，在过去两周约有 10,000 个被激活。[23]回顾土耳其选举，埃尔多安的主要政治对手克勒奇达尔奥卢指责俄罗斯使用深度伪造和其他人工智能生成的信息来干预土耳其选举，俄罗斯机器人账户、推特账户、各种虚假信息活动也支持埃尔多安，因为克勒奇达尔奥卢在民意调查中保持微弱领先。另一位候选人因杰声称自己是网上伪造的“性格暗杀”的受害者并退出竞选。《卫报》报道，因杰声称所谓性爱录像带是使用深度伪造技术制作的，使用了以色列色情网站的镜头，制造出看起来和听起来都像真实人物和事件的视频和图像。[24]土耳其通讯部长法赫雷丁·阿尔通（Fahrettin Altun）在推特上回应称，“我们打击虚假信息，我们认为虚假信息是民主的最大敌人之一，并准确、快速和透明地向国内和国际舆论提供信息。” [25]在土耳其选举前，克勒奇达尔奥卢就多次警告，竞选对手可能使用“剑桥分析式”策略和使用深度

伪造视频干预选举。[26]美国特别检察官罗伯特·穆勒（Robert Mueller）2019年的报告发现，俄特工试图影响美国2016年大选的结果，以支持特朗普。[27]战略与国际研究中心（CSIS）2020年的一份报告称，从2014年开始，俄罗斯开始瞄准东欧以外的国家，并扩大了策略，包括网络攻击和在线虚假信息活动，意大利、德国和西班牙的选举和政治进程都是目标。[28]2016年的美国总统大选和英国脱欧公投，剑桥分析公司利用 Facebook 用户数据画像选民心理，利用人工智能等技术对选民进行微定向宣传，帮助特朗普胜选和英国实现脱欧。那么人工智能、深度伪造对政治选举结果到底起到什么作用，产生多大影响？目前的可能结论是，对于实力相差悬殊的选举，人工智能、深度伪造或许能对选举产生局部性影响，在一城一地出现选情反转，但不足以产生全局性颠覆性影响。对于势均力敌的选举，人工智能、深度伪造通过局部反转，足以决定选举结果。例如，2016年的美国总统大选，剑桥分析公司帮助特朗普在个些州获得微弱优势，几万张选票反转就可以使选情产生大逆转。

2. 少量深伪内容能以独特方式破坏政治选举的信息空间。

自2022年11月 ChatGPT 发布以来，许多人关注人工智能通过生成虚假信息以颠覆政治选举的可能性。大量评论认为人工智能可能加速虚假信息生产，在2024年制造一场“虚假信息噩梦”。2023年，谷歌新闻索引结果到近3万篇新闻报道关注生成式人工智能

工具会如何影响选举。然而时至今日，生成式人工智能尚未像最初预期的那样改变信息格局。尽管人工智能生成内容在整个信息空间中仍只占一小部分，但即使规模较小，经人工智能完全生成或经过大幅修改的内容仍然可以通过各种方式破坏民主话语和选举公正性。在政治选举的大背景下，即使是少量的虚假生成内容也会破坏选举信息环境。例如，在选举最后关头伪造一段视频对候选人进行丑化或诬蔑，影响选民正常投票。2023年9月，生成式人工智能的政治干预对斯洛伐克议会选举产生颠覆性影响。此次议会选举事关斯洛伐克对乌克兰的军事援助以及对北约的支持。然而，就在选民投票的两天前，一段带有生成内容标记的音频在社交媒体大肆传播。该音频据称是亲北约的斯洛伐克进步党领袖米哈尔·希梅奇卡（Michal Šimečka）和一位记者讨论如何操纵选举的录音。尽管这段音频从一开始就显得很可疑，但还是在社交媒体上被数千人转发，其中包括斯洛伐克反对党的一名前国会议员。由于该音频存在声音不一致、措辞别扭、语句节奏可疑等异常现象，核查人员很快对录音的真实性提出质疑。然而，斯洛伐克的选举法要求媒体和政界人士在选举前的48小时保持沉默，避免发布与选举有关的公告，阻碍了纠正性信息的广泛传播。[29]这段伪造音频还利用了Meta的媒体操纵对策的缺陷，该对策明确规定只处理完全伪造的视频，而不处理音频。尽管核查人员最终在Meta上为该

音频贴上伪造标签，但由于不同平台审核标准不同，这段音频仍在碎片化信息空间广泛传播。这种最后一刻出现的丑闻，对于一些国家而言难以应对，因为这些国家限制媒体在选举前夕讨论与选举相关的内容，限制时长通常为 24 小时左右，有的可能长达三天，这显然会为揭穿网上疯传的虚假内容带来困难。

再如，声称某些真实信息是人工智能生成的从而逃避选民问责，尤其是针对某些可能影响政治竞选的真实丑闻。这个操作在阿根廷 2023 年总统选举中起到意想不到的作用。在第一轮投票前几天，网上开始广泛流传丑闻录音。据称，这些录音中时任总统候选人帕特里夏·布尔里奇（Patricia Bullrich）的经济部长人选卡洛斯·梅尔科尼安（Carlos Melconian）对女性恶语相向，并以政府职位换取性好处。这一事件被称为“梅尔科尼门”。事发后，布尔里奇和她的政党迅速为梅尔科尼安辩护，称录音是捏造的，可能是用“语音和人工智能”篡改或生成的。她还质疑音频片的来源，并攻击分享泄露音频的记者，指责他过去有不道德行为，并将篡改的音频作为支持现任者的抹黑活动的一部分。起初，梅尔科尼安选择对音频泄露保持沉默。不过，在随后的采访中，他并没有明确否认录音的真实性，而是说：“即便真是我说的，这又能说明什么呢？什么都没有。”至于这些音频片段是否真的是人工智能生成的深度伪造，尚无定论。不过，这

一事件凸显出，即使某些信息可能是人工智能生成的，它也能以意想不到的方式塑造选举竞争的轮廓。因为存在使人易信的深度伪造可能性，在未来，政客们将能够合理地将真实的丑闻斥之为捏造内容，即所谓的“骗子红利”。这种情况在美国已经出现，例如前总统唐纳德·特朗普的一段真实旧片段已经改头换面，成为人工智能生成的内容。有证据表明，这种策略对政客们很有效。最近的一项调查实验发现，把真实的丑闻标榜为“错误信息”会让选民更倾向于支持受牵连的政客。因此，生成式人工智能可能会在根本不使用它的地方产生最有害的影响。

[1] <http://www.yingyushijie.com/shop/source/detail/id/2505.html>.

[2] <https://www.euronews.com/next/2023/05/12/ai-content-deepfakes-meddling-in-turkey-elections-experts-warn-its-just-the-beginning>.

[3] <https://www.wired.com/story/deepfakes-cheapfakes-and-twitter-censorship-mar-turkeys-elections/>.

[4] <https://www.al-monitor.com/originals/2023/05/turkey-elections-deepfakes-disinformation-misdirect-voters-ahead-runoff>.

[5] <https://www.al-monitor.com/originals/2023/04/al-monitor-premise-poll-turkeys-election-dead-heat-erdogan-and-kilicdar>

oglu-tied.

[6] <https://coinmarketcap.com/community/articles/6465cd0713207b526cd69c67/>.

[7] <https://coinmarketcap.com/community/articles/6465cd0713207b526cd69c67/>.

[8] 一个被土耳其和美国指定为恐怖分子的组织。

[9] <https://www.wired.com/story/deepfakes-cheapfakes-and-twitter-censorship-mar-turkeys-elections/>.

[10] <https://www.intellinews.com/big-election-blow-for-erdogan-as-third-candidate-ince-abruptly-exits-contest-complaining-of-deepfake-sex-tape-278465/>.

[11] <https://www.theguardian.com/world/2023/may/11/mu-harrem-ince-turkish-presidential-candidate-withdraws-alleged-sex-tape>.

[12] <https://www.al-monitor.com/originals/2023/05/turkey-s-kilicdaroglu-accuses-russia-creating-deepfakes-against-opposition>.

[13] <https://www.timesofisrael.com/erdogans-main-rival-alleges-russia-posting-deep-fakes-in-lead-up-to-poll/>.

[14] <https://www.al-monitor.com/originals/2023/05/turkey-s-opposition-warns-against-deepfakes-ahead-fateful-elections>.

[15] <https://www.reuters.com/world/middle-east/erdogan-r>

ival-says-has-evidence-russias-online-campaign-ahead-turkey-vote-2023-05-12/.

[16] <https://www.fudzilla.com/news/ai/56884-deepfake-used-to-kill-of-turkish-presidential-candidate-s-campaign>.

[17] <https://www.al-monitor.com/originals/2023/05/turkey-elections-deepfakes-disinformation-misdirect-voters-ahead-runoff>.

[18] <https://www.al-monitor.com/originals/2023/05/turkey-elections-deepfakes-disinformation-misdirect-voters-ahead-runoff>.

[19] <https://www.dw.com/en/fact-check-turkeys-erdogan-s-hows-false-kilicdaroglu-video/a-65554034>.

[20] <https://www.newsweek.com/erdogan-rival-russia-turkey-election-1799962>.

[21] <https://www.dw.com/en/fact-check-turkeys-erdogan-s-hows-false-kilicdaroglu-video/a-65554034>.

[22] <https://www.telegraph.co.uk/news/2023/05/14/turkey-deepfake-elections-erdogan-muharrem-ince/>.

[23] <https://www.newsweek.com/erdogan-rival-russia-turkey-election-1799962>.

[24] <https://www.foxnews.com/world/deepfakes-porn-tapes-bots-ai-shaped-vital-nato-allys-presidential-election>.

[25] <https://www.intellinews.com/kilicdaroglu-warns-of-armed-elements-and-deepfake-videos-ahead-of-his-attempt-to-retain-erdogan-277952/>.

[26] <https://www.turkishminute.com/2023/05/08/erdogan-play-deepfake-video-at-election-rally-to-show-opposition-linked-to-terrorist-group/>.

[27] <https://www.newsweek.com/mueller-report-donald-trump-russia-1400935>.

[28] <https://www.csis.org/blogs/strategic-technologies-blog/russia-ramps-global-elections-interference-lessons-united-states>.

[29] <https://www.brookings.edu/articles/the-impact-of-generative-ai-in-a-global-election-year/>.