



2024

人工智能赋能电子数据分析 的机遇与挑战

吴松洋

公安部第三研究所

人工智能大模型

机器学习

分类

回归

聚类

关联规则

深度学习

CNN

RNN

大模型/生成式

AI

大语言模型

多模态模型

人工智能 (AI)

是一种模拟人类智能的技术。它涵盖了机器学习、深度学习、自然语言处理等一系列专有技术，使计算机能像人一样思考、学习和解决问题

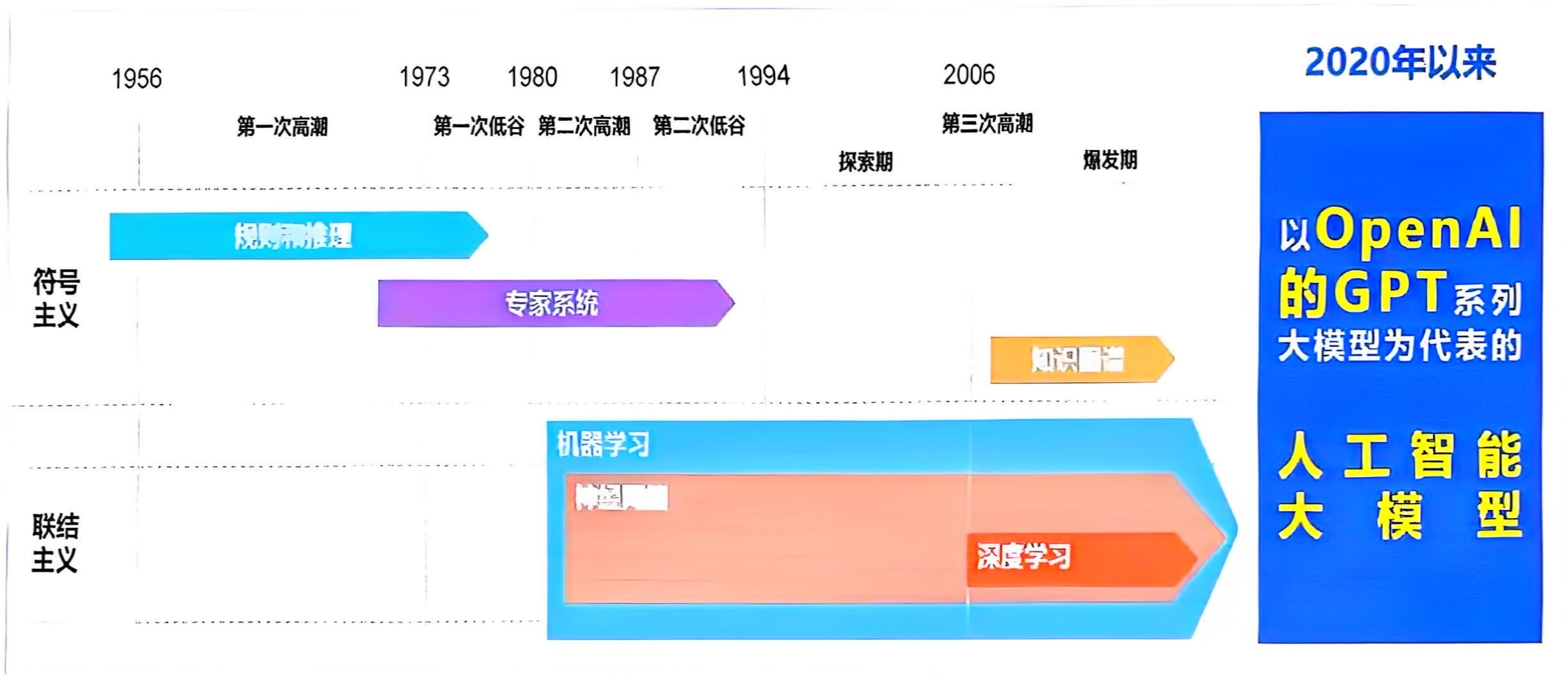
狭义人工智能 (ANI)

是经过训练并专注于执行特定任务的人工智能。当前，我们周围的大部分人工智能都属于狭义人工智能，如语音助手、自动驾驶汽车等

通用人工智能 (AGI)

是人工智能的一种理论形式，机器将具有与人类相同的智能；它会有自我意识，有能力解决问题、学习和规划未来

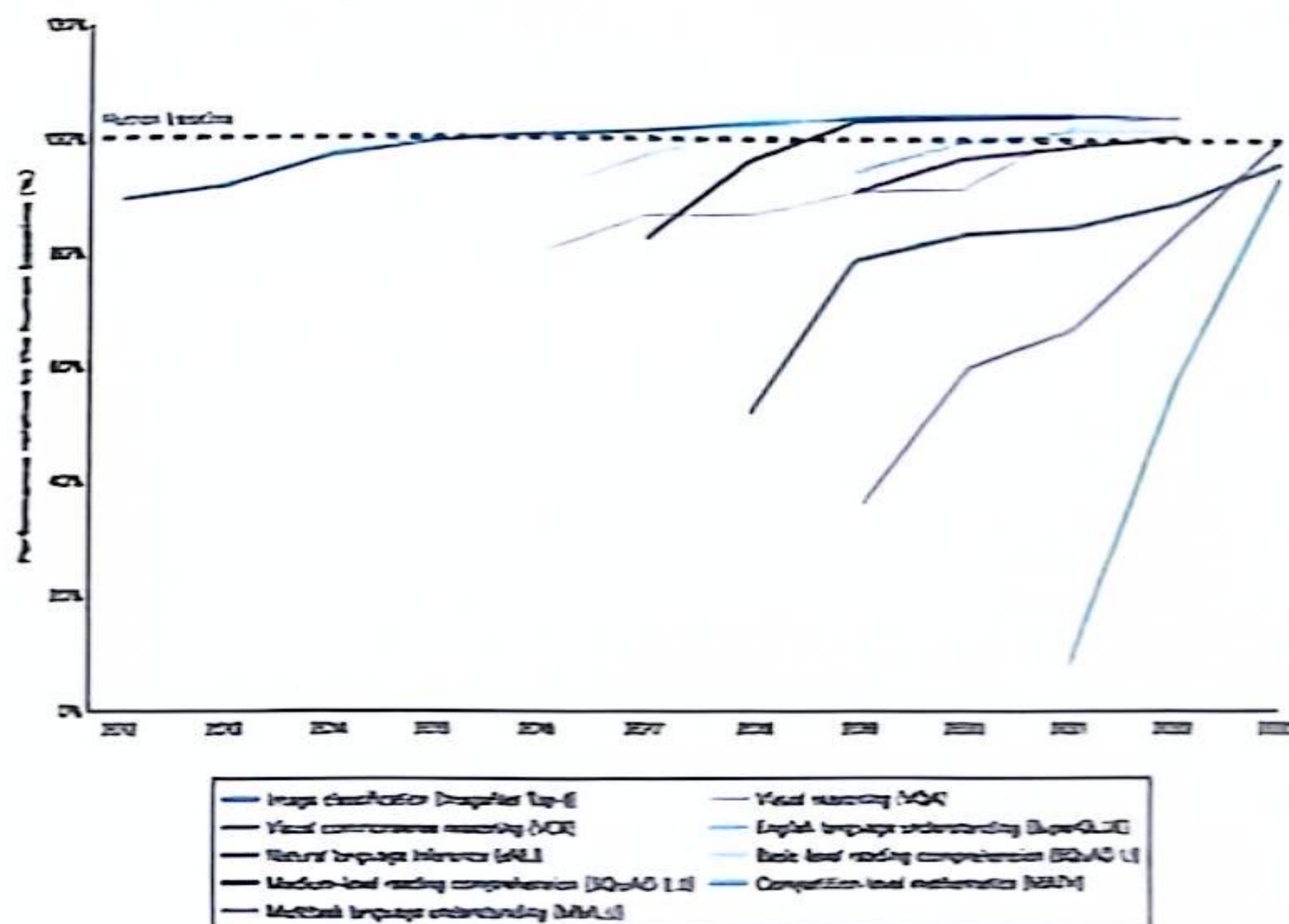
人工智能大模型



人工智能是一项由机器代替人类实现认知、识别、分析、决策等功能的技术，其本质是对人类意识与思维信息过程的模拟

人工智能大模型

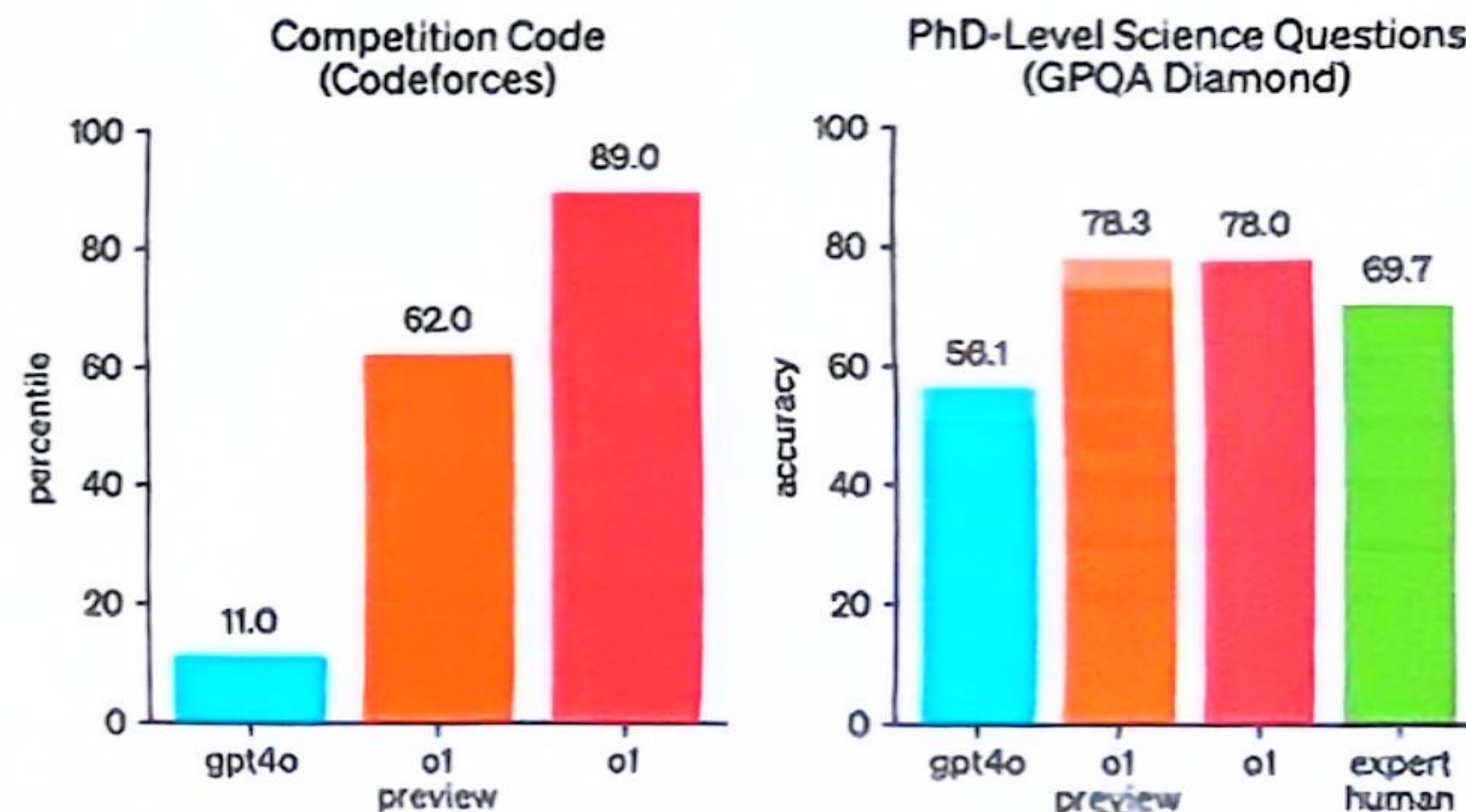
Select AI Index technical performance benchmarks vs. human performance
Source: AI Index, 2024 / Chart: EY AI Index report



截至 **2023 年底** 人工智能在大部分任务中的表现已经超越了人类的能力，目前人类只有在复杂的认知任务还有一点优势

—— AI Index 2024 <https://aiindex.stanford.edu/report>

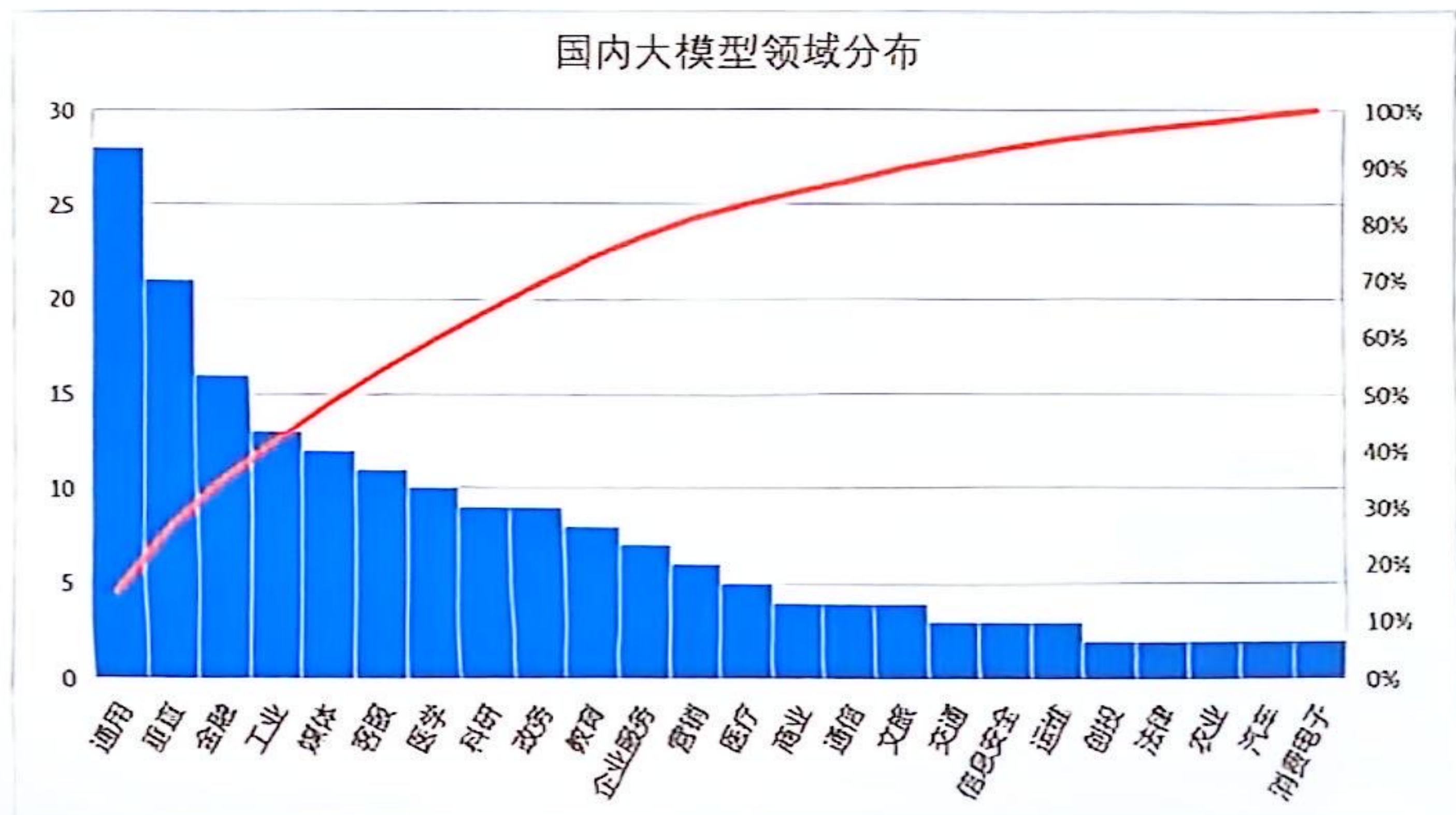
2024年9月13日 OpenAI发布o1系列模型，在博士级科学推理问题测试中，o1模型表现出了接近甚至超越人类顶尖水平的实力



国内大模型技术现状

国内大模型厂商主要包括百度、阿里、讯飞、华为等企业，也有智源研究院、中科院自动化所等研究机构，目前公开报告可见的厂商近300家，发布模型超过400种

——<https://github.com/wgwwang/LLMs-In-China>



- 大多均采用MaaS模式
 - LLaMA 3.1、QWen 2.5 等
- 开源模型能力不断提升

人工智能大模型

人工智能

自然语言处理

自动标签

数据挖掘

话题聚类

语义分析

情感分析

辅助侦查

计算机听觉

语音识别

自动翻译

声纹鉴定

计算机视觉

人脸识别

OCR识别

图像视频鉴定

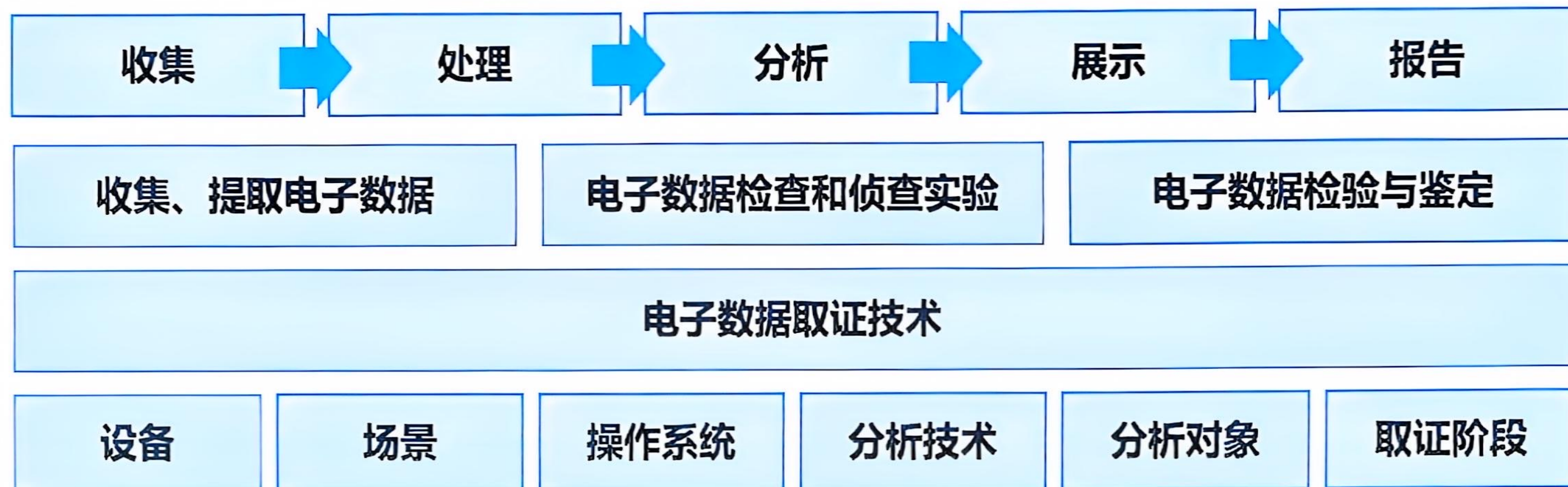
结构化数据处理

数据检索

证据展示

案件知识库

电子数据取证分析



- | | | | | | |
|-----------|----------|---------------|----------|-----------|------|
| • PC机取证 | • 现场取证 | • Windows取 | • 网络数据分析 | • 注册表 | • 发现 |
| • 手机取证 | • 网络在线提取 | • Linux取证 | • 密码破解技术 | • 日志 | • 固定 |
| • IOT设备取证 | • 远程勘验 | • Android取证 | • 数据恢复技术 | • 文件 | • 提取 |
| • 无人机取证 | • 侦查实验 | • Mac OS取证 | | • 数据库 | • 分析 |
| • 汽车取证 | • 检验鉴定 | • Apple IOS取证 | | • 病毒和木马分析 | • 展示 |

电子数据取证分析

电子数据取证分析业务工作

海量数据

异构数据

加密数据

数据分析

推理归纳

证据展示

反取证分析技术

新兴技术

痛点、难点

处理效率
难提升

证据分析
难深入

业务场景
变化快

人工智能大模型技术，解决电子数据取证分析业务中的痛点、难点

PART 02

AI对电子数据取证的 机遇



对电子数据取证分析的赋能



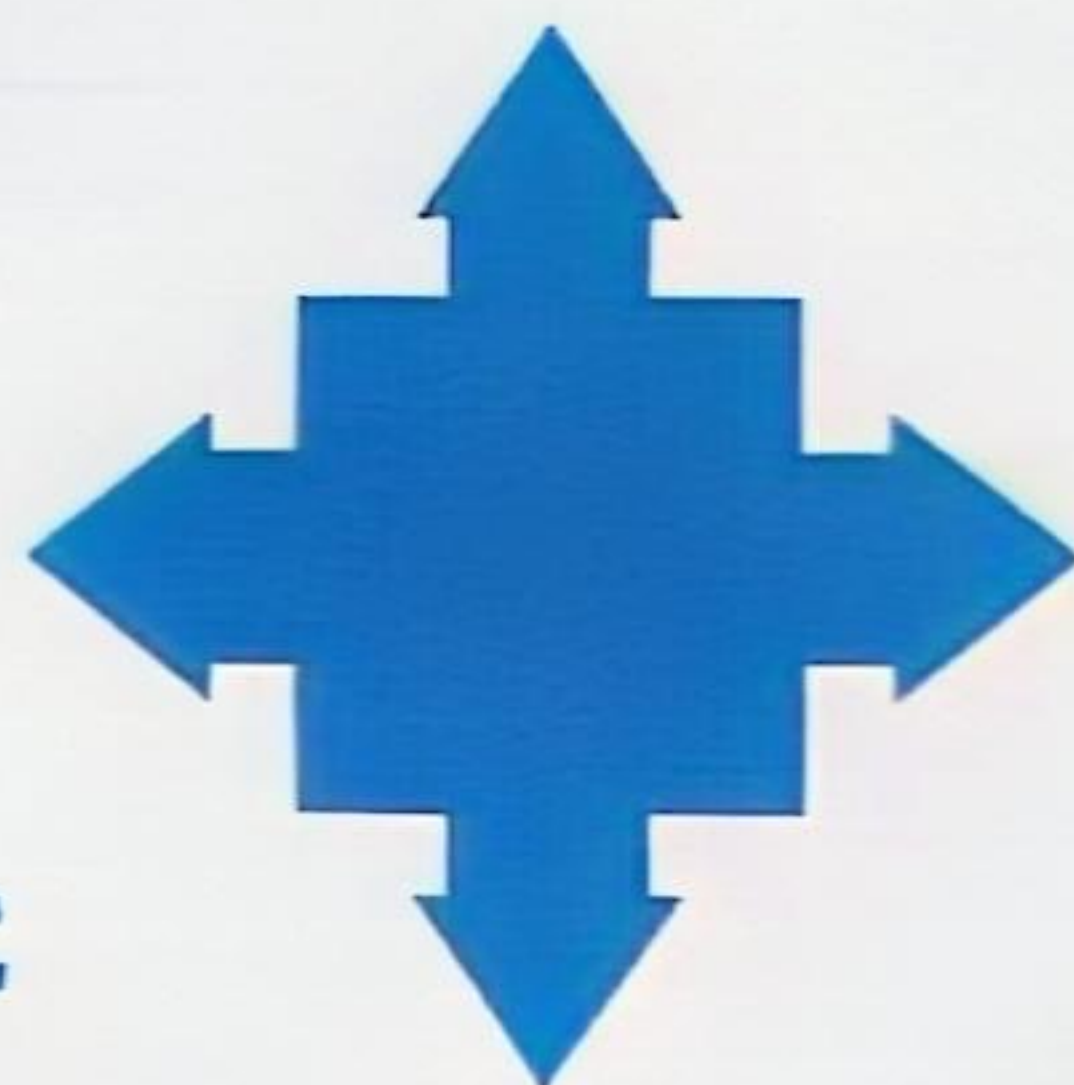
强化电子物证收集

- 聊天记录、社交账号、购物信息等重要电子证据规模庞大，且往往会分散存储在物联网、智能手机等多个终端或云平台上
- 人工智能能够助力，公安机关快速收集、整合大量数据，帮助缩小审查范围



提高现场勘验效率

- 智能化无人装备现场勘验技术，可以捕捉到侦查人员不易察觉或侦查难度较高的角落和细节
- 精准获取被伪装或隐藏的现场线索和证据，让现场证据的提取更科学、准确。



案件线索研判分析

- 海量线索和证据中，自动识别案件特征，确定犯罪嫌疑人的犯罪模式
- 可为并案侦查和案件研判提供辅助决策支持



辅助证据审查

- 对证据的合法性和取证的操作程序进行把关和审查，确保证据的效力
- 根据人工智能给予的证据收集指引，全面地收集证据，防止在证据链条环节出现遗漏

对电子数据取证分析的赋能

场景一 文本挖掘

1. 基于关键词规则的线索发现

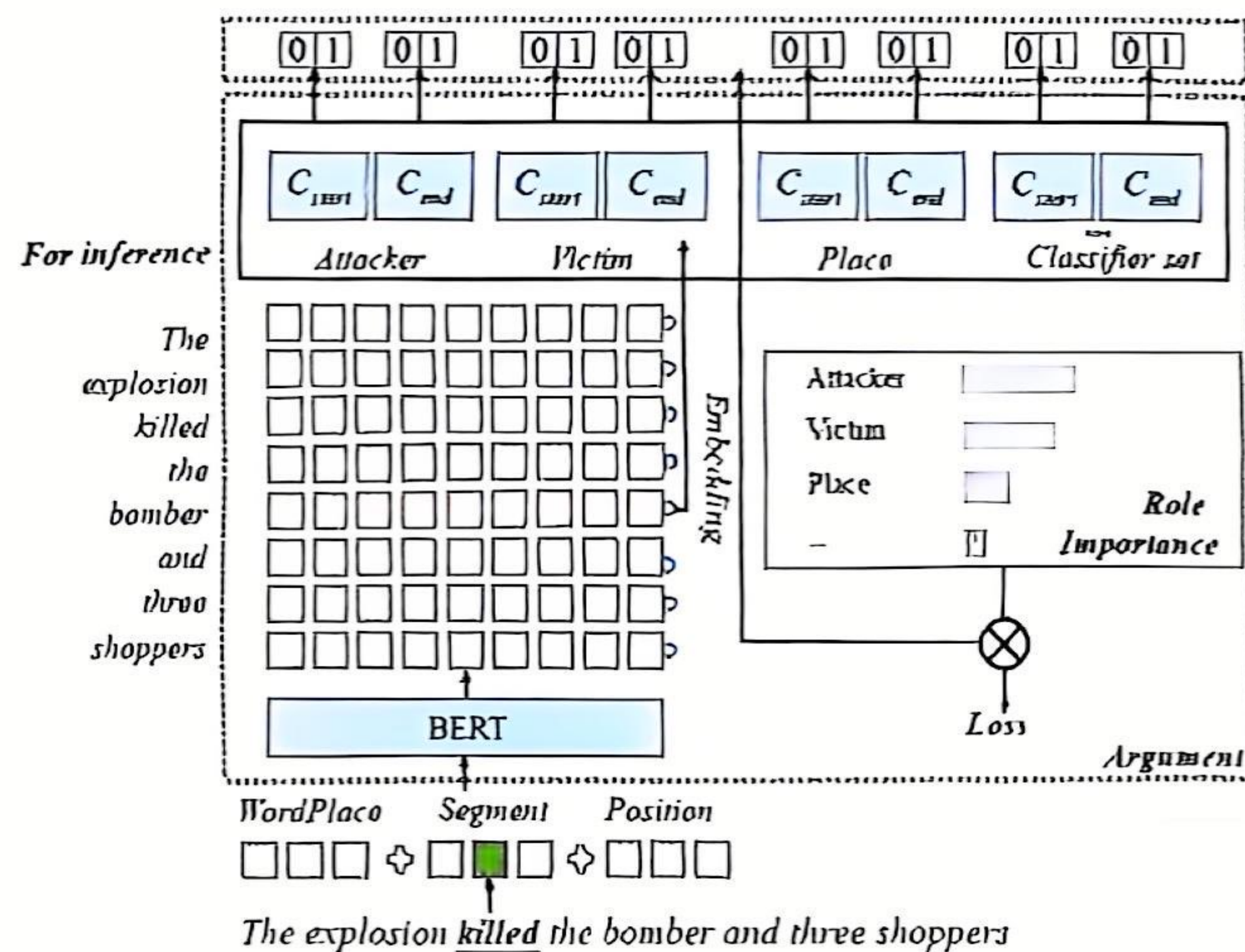
黄、赌、毒、诈骗、传销等案件中，相同案件类型的电子数据文本内容中会出现相似的文本关键词、代号、暗语、聊天会话模式和网址信息等

2. 基于AI模型的线索特征分析

通过对类案数据抽取、标注预处理等，构建训练人工智能文本分析模型，学习不同案件文本表达语义特征，发现特定案件类型的文本线索

3. 复现办案民警分析思路

有经验的办案民警往往能够从犯罪分子的微信、微博等应用的文本中快速发现案件线索



对电子数据取证分析的赋能

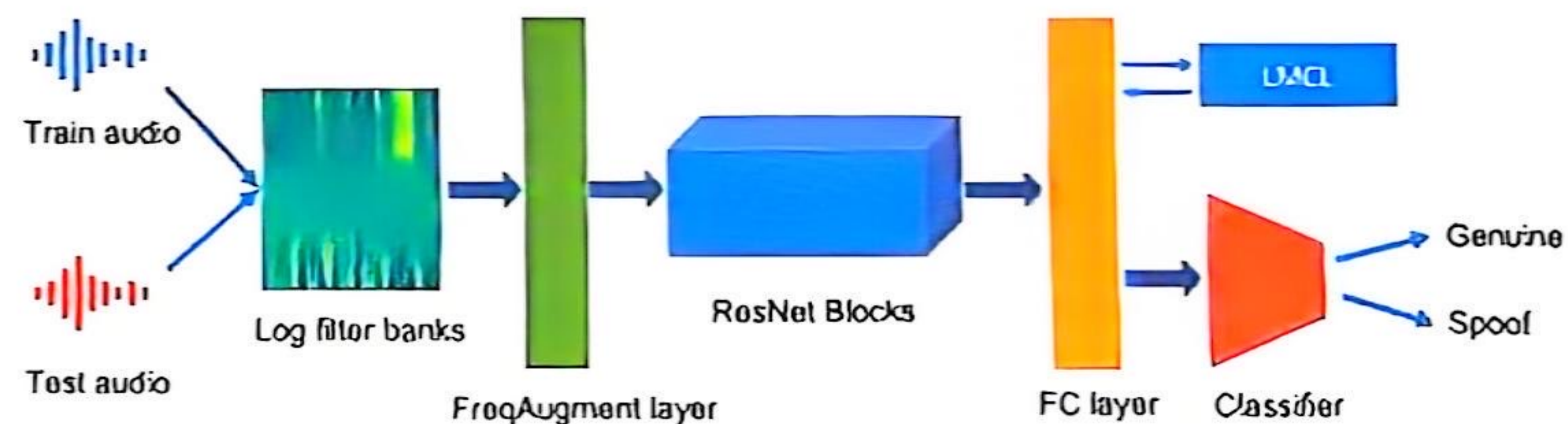
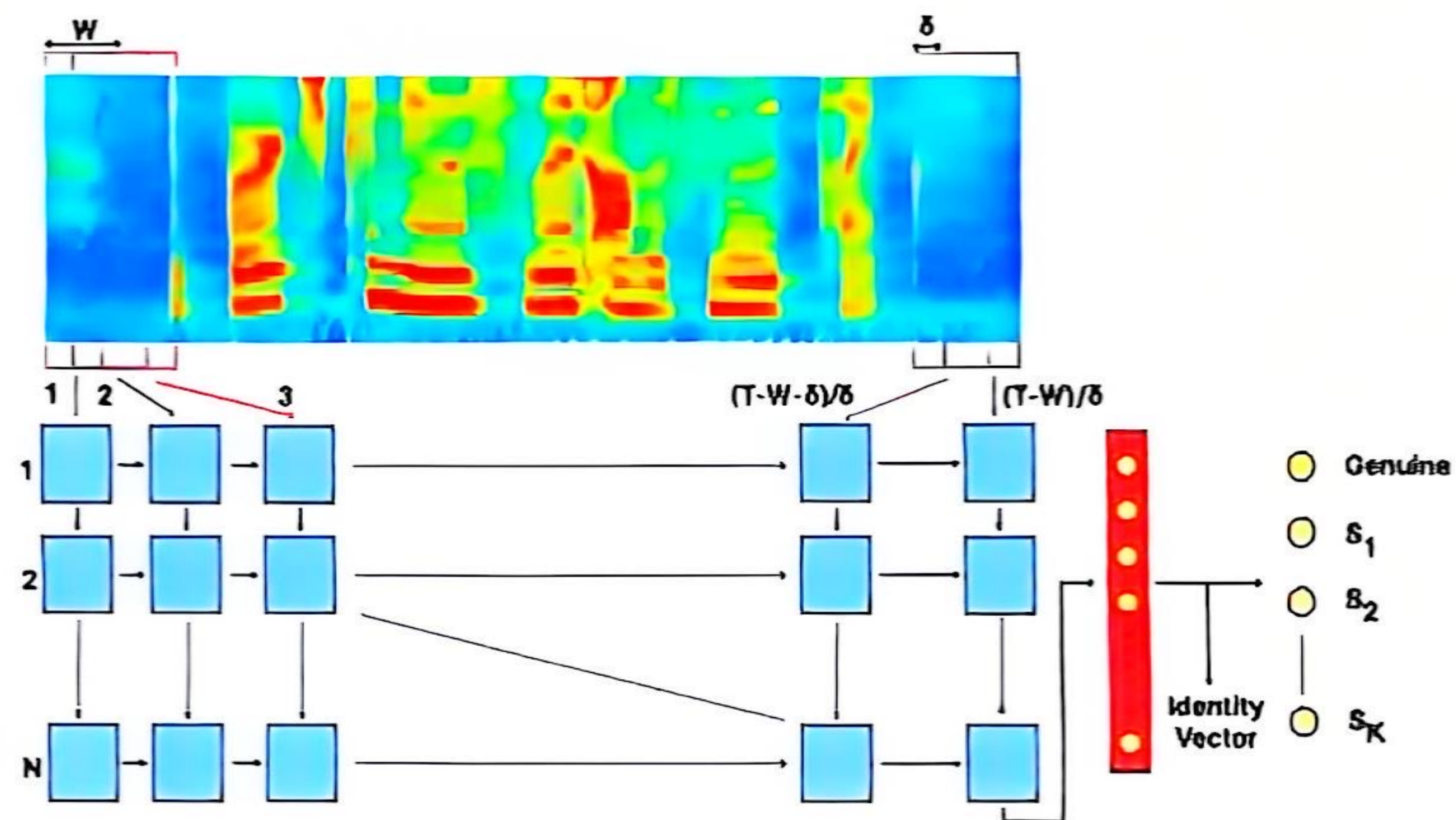
场景二 语音分析

1. 语音转文字

利用人工智能技术训练语音识别模型，自动批量分析语音数据，将其转换为文本数据，发现识别涉案线索，提高案件数据分析的效率

2. 声纹识别与鉴定

每个人的声音特征是一种特有的生物信息特征，利用人工智能技术建立广泛的声纹库，使用图像识别技术和声纹比对技术实现对电子数据中声纹的快速识别和比对，快速确定案件线索



对电子数据取证分析的赋能

场景三 图像分析

1. 人脸识别

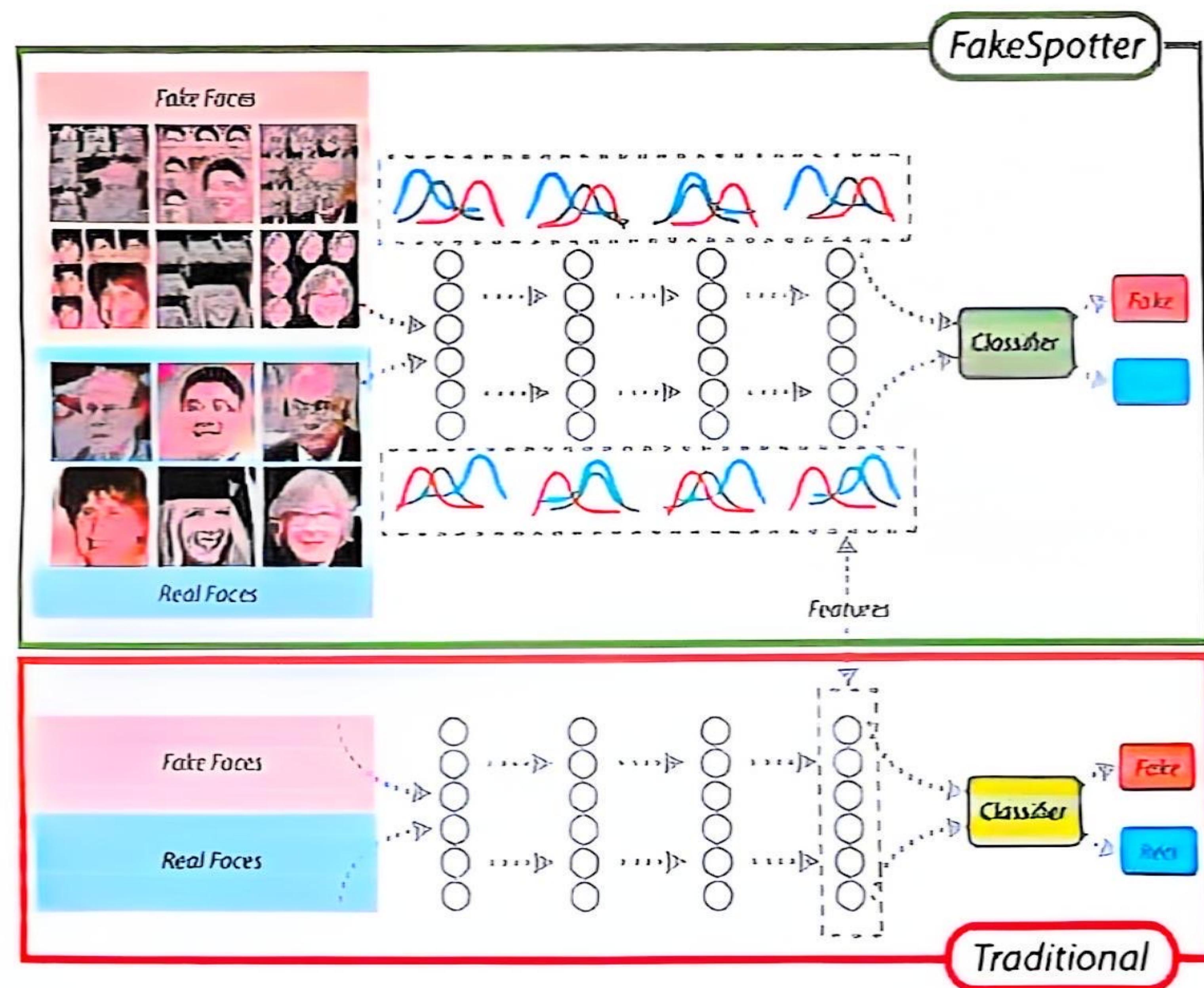
利用人脸识别、人脸检索、人脸比对等技术可实现对电子数据中特定人物人脸的发现、比对和检索

2. 影像内容识别

对各类证照、银行卡、交易单据、发票以及二维码等各类关键线索图片的识别，从海量电子数据中快速高效筛选出与案件相关的重要线索

3. 影像伪造鉴别

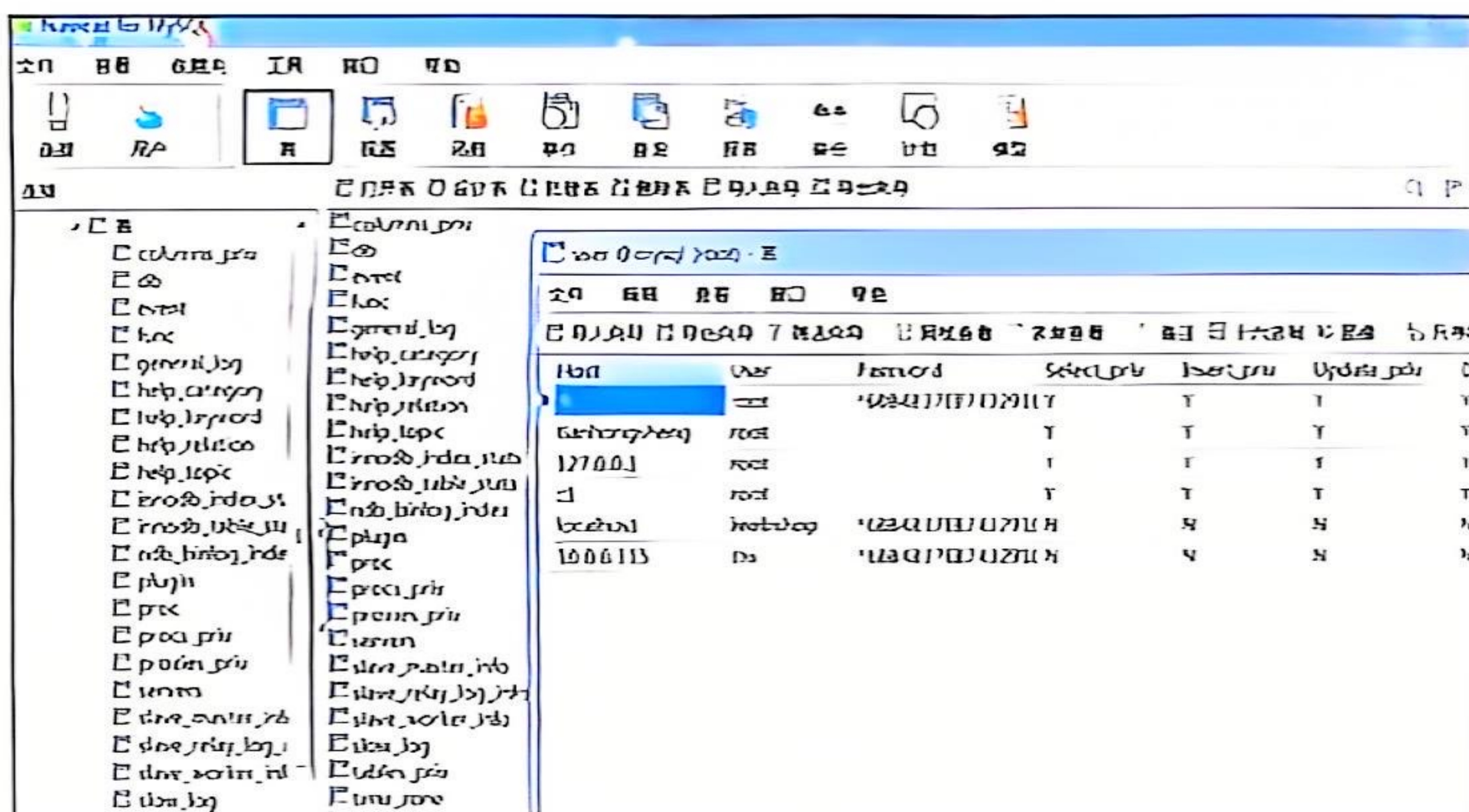
一些不法分子利用这类技术伪造虚假照片和视频，严重危害受害人的个人名誉，可利用人工智能技术训练伪造图像视频识别模型



人工智能赋能取证数据分析

取证数据库资金智能分析

取证数据库作为案件扩线分析重要数据，蕴含着重要线索，且价值密度高。但由于数据库种类多样，数据表数量巨大，字段名无规则，传统技术很难从中快速分析出有用的案件信息，需要大量手工操作



某四方支付平台后台数据库fourpay, 中有
79张表, 通过清除空表后, 还剩**55张**
表, 需要民警逐个打开识别

人工智能赋能取证数据分析

数据表智能识别与推荐

- ✓ 基于实际案件取证分析需求;
- ✓ 取证数据作为案件关键扩线分析, 现状大多时候人工分析;
- ✓ 数据库种类多样, 数据表数量巨大, 字段名无规则;
- ✓ 实际案件分析过程重点关注人员、资金交易等信息

数据名称	数据类型	数据大小	数据表的数量	案件定性
mpzy	Access数据库	8.62MB	159	涉黄
小圈	Mysql数据库	89.6GB	113	涉黄
108world	Mysql数据库	1.44GB	50	跑分
gallant	Mysql数据库	13.4MB	64	跑分

- moment.sql
- operator_special_log.sql
- option.sql
- pay_channel.sql
- pay_data.sql
- prize.sql
- region.sql
- reply.sql
- reply_data.sql
- report.sql
- report_feedback.sql
- resource.sql
- resource_check.sql
- role.sql
- stage_record.sql
- ums.sql
- ums_ext.sql
- spam_words.sql
- stat_daily_active.sql
- stat_daily_appname.sql
- stat_daily_consume.sql
- stat_daily_limit.sql
- stat_daily_register.sql
- stat_daily_new_user.sql
- stat_daily_recharge.sql
- stat_daily_trade.sql
- stat_month_login_log.sql
- stat_user.sql
- mailcn.sql
- log.sql
- task.sql
- task_prize.sql
- topic.sql
- trade.sql
- trade_balance.sql
- trade_buy.sql
- trade_income.sql

[illegible]

人工智能赋能违法网站分类

违法网站分类识别

- ✓ 公安民警日常巡检过程中，需要识别研判违法网站
- ✓ 根据标题、时间、倾向性、用户、摘要、关键词、链接等结构化文本知识，网站截图、网站关系图等信息对网站进行分类识别
- ✓ 网站数量众多，人工巡查或传统人工智能效能较弱

违法网站种类

网络赌博

刷单类诈骗

色情

贷款类诈骗

网络攻击

裸聊诈骗

投资理财诈骗

四方支付

门户网站

其他

人工智能赋能违法网站分类

基于大模型的 违法网站分类识别

- ✓ 使用**GPT4 API**进行零样本识别
 - ✓ 面向网页代码的知识抽取
 - ✓ 面向截图的网站分类
 - ✓ 准确率**超过93%**
- ✓ 使用本地**LLAMA2-7B(16bit)**进行零样本识别
 - ✓ 需要**2块T4显卡 (32G)**
 - ✓ 准确率达到**80%**



PART 03

AI对电子数据取证的挑战



对电子数据取证分析的挑战



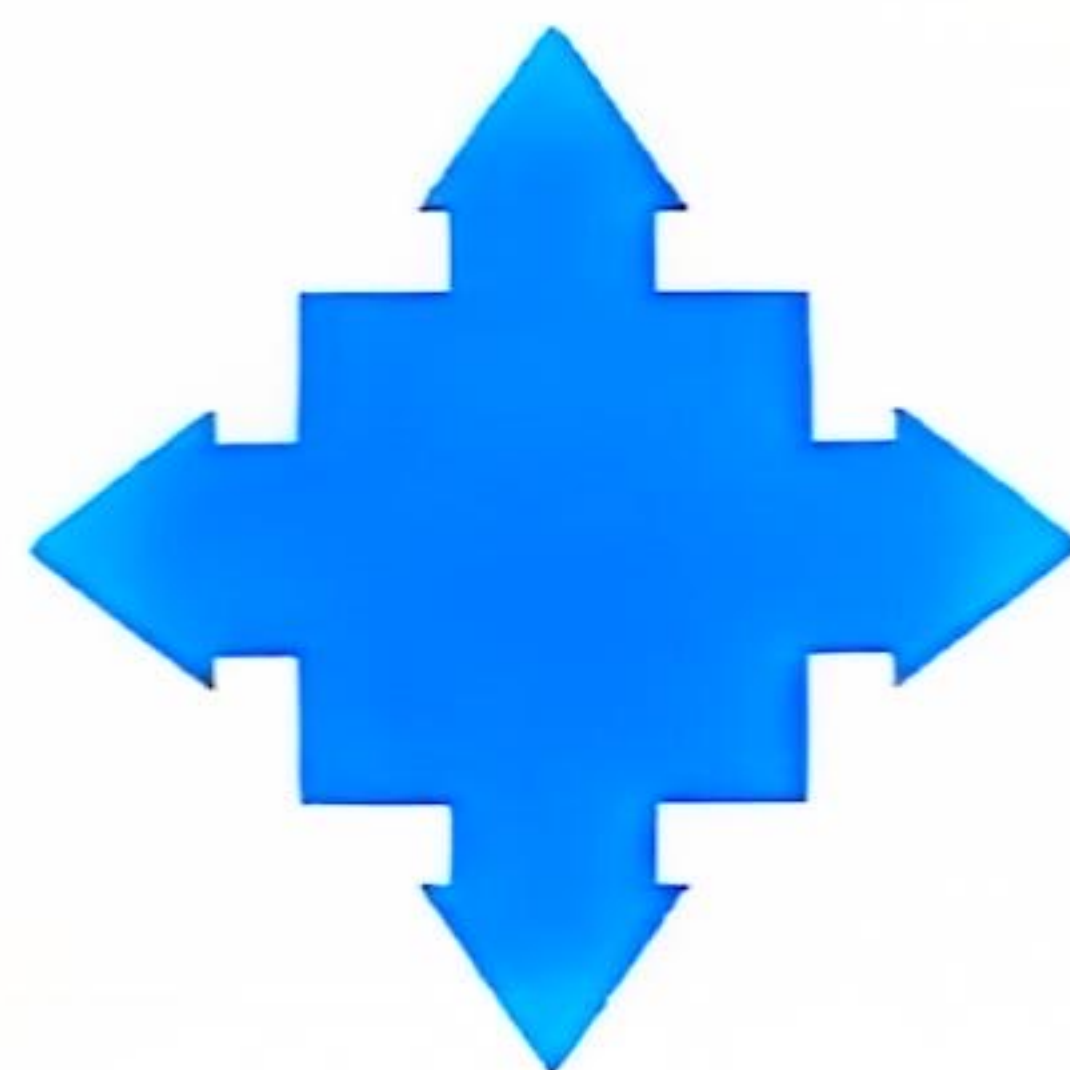
不法分子反侦查能力不断提高

- 人工智能技术被频繁用于新型网络犯罪活动及反侦查上，如AI聊天机器人被用于黑产引流，AI换脸被用于电信诈骗等等，犯罪隐蔽性更强



人工智能输出幻觉易出现疏漏

- 人工智能在实际应用过程中可能表现为纯粹的机器行为，完全按照预定的模式执行，还会存在算法不成熟、输出幻觉内容、数据不完备等问题，导致人工智能在一些复杂情况下的取证和决策存在不确定性



人工智能的算法可能存在偏见

- 人工智能会受到预训练模型、自身算法生成逻辑及背后公司的研发人员影响，可能会无意识地使算法呈现出一定的偏向性



侵犯公民的隐私权

- 人工智能技术拥有强大算力能够快速收集互联网中数据库的海量信息，但即使是互联网中公开的公民信息也有可能涉嫌侵犯公民的个人权利

对电子数据取证分析的挑战

生成式人工智能技术引发多媒体内容安全隐患

特朗普虚假视频



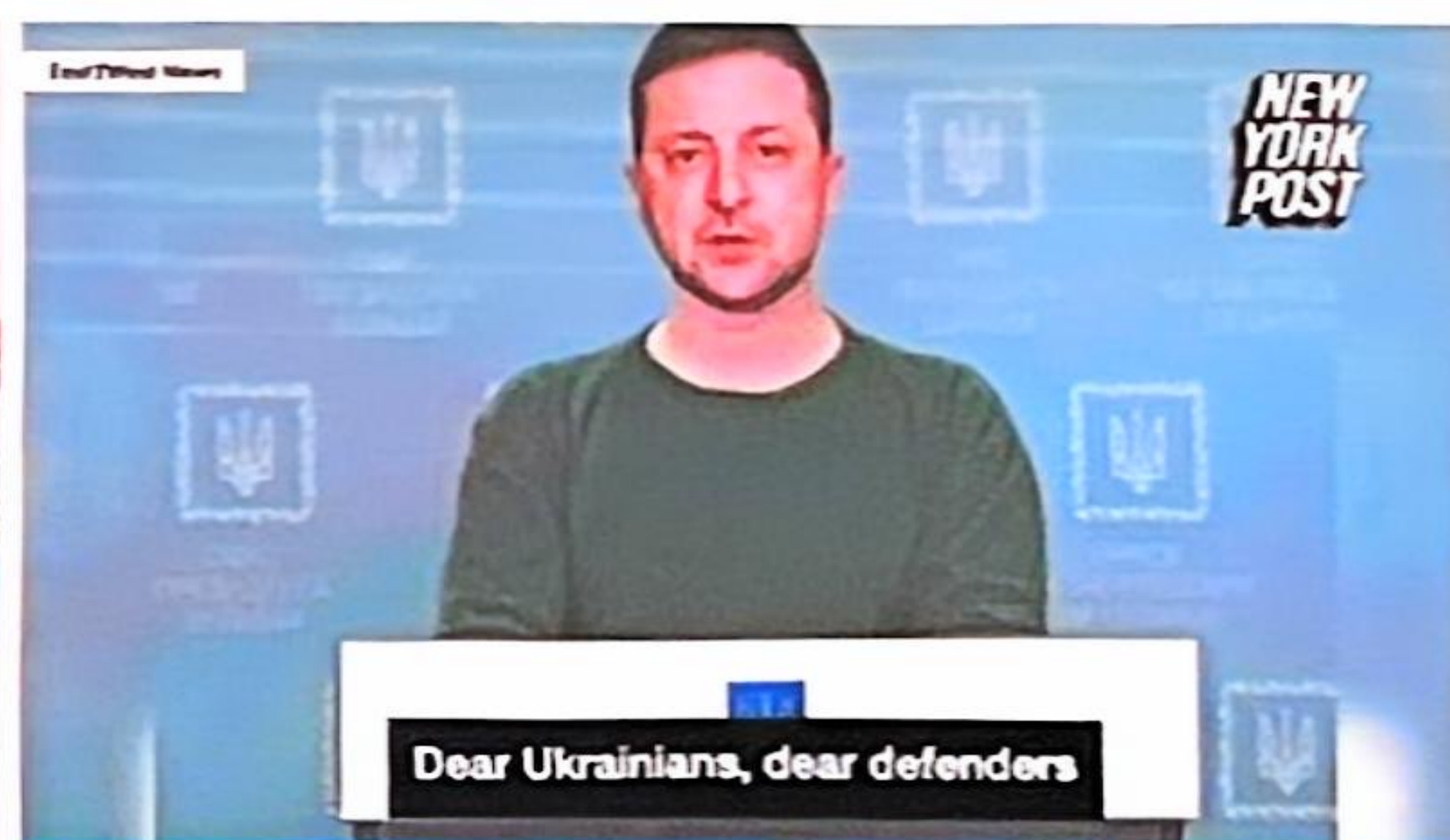
侵害国家政治安全

针对企业电信网络诈骗



侵害人民财产安全

乌克兰总统投降视频



战时投放 影响军心

恶意使用生成式人工智能技术，所生成的音视频能够被用于**诈骗**、甚至影响**国家政治**，显示出了巨大的破坏力，将会给个人、企业造成了声誉损害和财产损失，乃至给社会、**国家安全**造成威胁

对电子数据取证分析的挑战



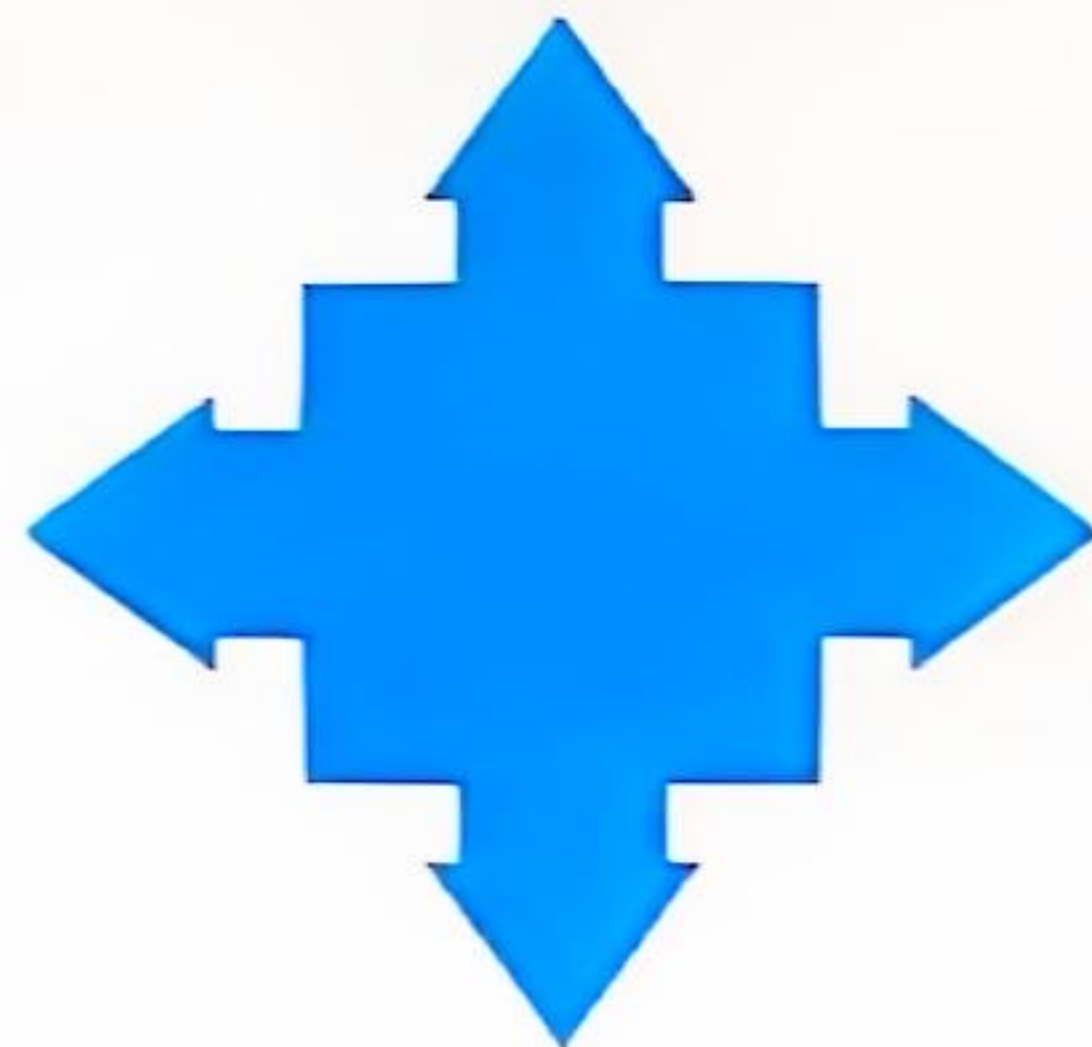
更直观电子数据分析

- 在人工智能技术的帮助下，高效完成跨语种文本翻译、视频转文字、图片转文字、语音转文字、以字搜图、以文搜图、以图搜图、图片分类和文档分类等任务



更高效社交应用线索挖掘

- 人工智能大模型可进行文本分析，快速识别和深度分析成员关系、成员角色、组织架构、话题及暗语等方面，并输出分析结果，提升证据链固定效率



更灵活电子数据线索扩线

- 依托人工智能大模型的多维分析能力，根据取证需求，灵活调整分析策略，实现角色、异常行为、社会关系等个性化多维分析挖掘及重点信息总结



更全面线索挖掘及报告撰写

- 依托人工智能大模型强大的泛化能力、复杂的语义及上下文理解能力，实现提供少量样本，即可精准识别和提取全局数据中的同类实体，输出可读性强、精准度高的总结报告