网络空间安全导论

主讲人： 杜瑞颖



目录

1. 信息内容安全的概念

信息内容安全，简称内容安全。它是信息安全在**政治、法律、道德**层次上的要求。

信息内容是安全的，就是要求信息内容在政治上是健康的， 在法律上是符合国家法律法规的，在道德上是符合中华民族优良的道德规范的。

因此，信息内容安全是信息在语义层次上的安全。

2. 信息内容的产生源

社交媒体(Social Media)指互联网上基于用户关系的信息内容生产与交换平台。是人们彼此之间用来分享意见、见解、经验和观点的工具和平台，现阶段主要包括社交网站、微博、微信、博客、论坛、播客、网络直播、搜索引擎等等。

社交媒体在互联网的沃土上蓬勃发展，爆发出令人眩目的能量，其传播的信息已成为人们浏览互联网的重要内容，不仅制造了人们社交生活中争相讨论的一个又一个热门话题，更进而吸引传统媒体争相跟进。

- 全球最大的视频网站一天上传的影像，可以连续播放98年。
- 今天，两天积累的信息总和就相当于人类历史留下的全部记忆

2. 信息内容产生源

基于大规模数据和云计算算力，人类将从逐渐重视算法结果的建议愈发深入转向直接接受算法结果的指令。通过机器学习、计算机视觉处理、语义分析与推理、自然语言分析交互等人工智能技术，算法日臻自学习化、智能化，在互联网**信息内容源生产、信息内容分发路径和信息内容生态**三大方面带来重大变革影响。

算法生成UGC（user generated content）成为互联网新的信息内容来源。集中体现在大型内容平台生产机制方面，信息内容生产速度、效率和质量都极大提升。一方面克服用户原创信息内容水平参差不齐、难以进行人工审核、缺少持久创作机制等问题；另一方面伴随人工智能和物联网纵深普及，信息内容生成算法将在海量丰富、源源不断的非结构化数据训练下更加精益智能。此外，信息内容生成算法充分挖掘整合了此前人类经验积累的标准化模板，并以此为基础进一步优化完善。

2. 信息内容被操控（攻击）

算法和人工智能技术也带来内容生成被操纵控制的风险隐患。

首先是UGC 信息内容生产在智能化的同时更有欺骗性。如华盛顿大学的研究人员用神经网络来观看视频，并将不同的音频声音转换为基本的嘴形，根据音频剪辑生成逼真的视频，让嘴巴的运动和音频同步。

同样的技术原理可以制作一系列色情、仇恨、暴恐等视频内容，并将原视频的人物“换脸”。

其次是人工智能技术支持的语音转文字存在被恶意操纵的风险。近日谷歌研发成功Tacotron2系统，是文字语言转换TTS（Text-to-Speech）系统的一种，特点是不需要使用复杂的语言和声学特征作为输入，而是仅使用语音示例和相应文本记录训练的神经网络，从文本中生成类人语言。Tacotron2 系统目前主要的应用主体是谷歌智能语音助手。

目前，网络超声波攻击技术可利用语音助手麦克风的硬件漏洞，使语音助手接收并执行超声波指令，实现对语音助手设备的控制。语音助手一旦被恶意控制，则实施网络犯罪欺诈活动，或增加发布传播违法有害信息的风险。

2. 信息内容被操控（攻击）

信息隐匿话。加密算法、加密协议的持续演进，以及下一代互联网、IPv6 为代表的新技术普遍提供了匿名通信功能, 加密邮件, 加密即时通信等普遍应用, 网络行为的隐匿性越来越强。

加强用户保护通信隐私的同时, 也为互联网内容监测审查带来难度，给违法有害信息传播和网络欺诈犯罪以可乘之机。

互联网的开放性叠加隐匿性为加剧不良信息内容舆论传播扩散，逃避管控处置和追踪溯源推波助澜，为网络违法犯罪、极端暴恐、民族分裂仇恨活动提供了化整为零、隐蔽勾连的重要依托。

Twitter、Facebook、Youtube、微信 等社交媒体平台作为信息内容聚合交互的中心节点，主导着全球舆论生态的生成分发环节，结合其他加密社交软件、虚拟专用网、阅后即焚应用软件等网络通讯工具，通过社交分发和定向智能推荐分发模式，实现跨地区跨国界舆情生态的影响干预，并支撑起网络间谍、暴恐分裂和欺诈犯罪活动的行动勾联和信息沟通。

3. 信息内容的分析技术

文本挖掘，也称为文本数据挖掘，其主要目的是采用数据挖掘的技术，从非结构化或半结构化的语言文本中提取出潜在有价值的、新颖的、可被理解的、重要的模式和知识，是保障互联网信息内容安全的重要手段。

自然语言理解/处理(N LP，Natural Language Processing)是使用自然语言同计算机进行通讯的技术, 因为处理自然语言的关键是要让计算机“理解”自然语言,所以自然语言处理又叫做自然语言理解(NLU ,Natural Language Understanding), 也称为计算语言学(Computational Ling uistics)。一方面它是语言信息处理的一个分支，另一方面它是人工智能(AI，Artificial Intelligence)的核心课题之一 。

让计算机理解中文就是中文信息处理。

3. 信息内容的分析技术

机器学习，是实现人工智能的一种方法。机器学习的概念来自早期的人工智能研究者，已经研究出的算法包括决策树学习、归纳逻辑编程、增强学习和贝叶斯网络等。

简单来说，机器学习就是使用算法分析数据，从中学习并做出推断或预测。与传统的使用特定指令集手写软件不同，它使用大量数据和算法“训练”机器，由此带来机器学习如何完成任务。

人工智能，它是研究开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新型技术科学。该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。

从事这项工作的人必须懂得计算机知识，心理学和哲学。人工智能是包括十分广泛的科学，它由不同的领域组成，如机器学习，计算机视觉等。总的说来，人工智能研究的一个主要目标是使机器能够胜任一些通常需要人类智能才能完成的复杂工作。

4. 舆情控制

区块链技术

区块链的数据验证、共识更新、传输存储加密和智能合约自动执行等功能，成为防数据篡改，留存行为痕迹，创造匿名互信生态的关键技术支撑。

目前区块链在信息内容生产、分发领域的主要应用较为点状零星，有通过区块链的数据共识更新、智能合约执行机制来创建一项信息内容创作众包项目，实现多节点协作和精准计酬；有通过传输存储端到端加密、不可篡改来确保数据互验和信息分发的精准性和可溯源，创建真实防伪可追溯的信息内容生态，并实现多极各节点快速流转分发。

4. 舆情控制

大数据处理技术

利用大数据处理技术建立舆情监测预警平台，强化对重大敏感时间节点的舆情跟踪，深入开展内容生态的态势研判和规律分析；针对违法不良信息扩散、敏感热点舆情、暴恐极端活动事件的线索表征做到及时发现，积极疏导，应急响应；提升舆情正向干预和危机处置技术能力。

4. 舆情控制

管理措施

一、健全完善内容安全治理规范体系，明确细化电信运营商、互联网接入服务平台企业、互联网内容服务平台企业的信息内容安全保障责任；从内容审查过滤、用户实名管理、应急响应处置、日志留存和行为审计溯源、第三方审查评估、执法协助配合、举报投诉机制、合作方督促、平台自律执行和平台替代责任等方面强化过程与结果约束。

二、建立推动国家信息内容安全的跨部门情报共享、监测巡查、日常联动、应急处置、通报约谈、协同查处等机制运行和执行实施。

三、建立准入评估、问题巡查、执法检查、信用公示等监管措施合力，加强平台企业主体责任督导。

严格准入阶段的主体能力审查和新技术新应用的事前事中评估审查；依托安全评估巡查监测机制，组织开展责任落实专项督导，持续跟踪锁定问题隐患、督促整改，加强行政处罚和监督问责；研究建立重点平台互联网企业联席机制和责任信誉评价机制，将主流媒体平台和内容平台打造为公众与政府沟通交流的纽带和政府信息传播发声主阵地，以政府权威公信力提升带动正向舆论。

思考题

1. 信息内容安全对国家的重要性？

2. 简述人工智能技术对信息内容安全的影响。

谢谢！