

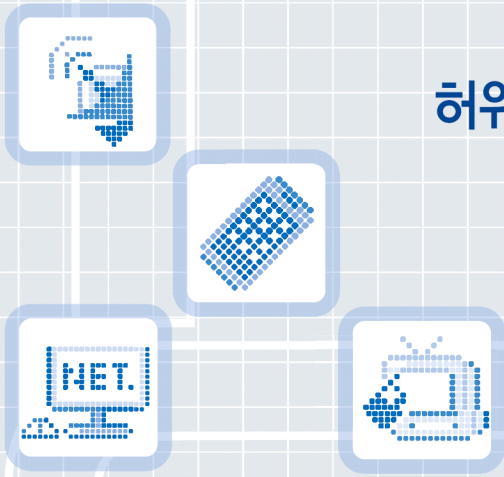
KISDI

Premium Report

허위정보 시대의 선거 개입과 인공지능

최계영

정보통신정책연구원 선임연구위원



Premium Report

허위정보 시대의 선거 개입과 인공지능

최 계 영

정보통신정책연구원 선임연구위원

요약문

- 1. 정보공간에서의 영향 공작(Influence Operation)과 선거개입 4
- 2. 인공지능 혁신과 위협 9
- 3. 대응 방향과 시사점 12
- 참고문헌 18

허위정보 시대의 선거 개입과 인공지능

최 계 영

정보통신정책연구원 선임연구위원

*choigi@kisdire.kr, 043-531-4321

*현 정보통신정책연구원

디지털경제사회연구본부 디지털경제연구실

요약문

본 리포트는 거대언어모델 등 생성 인공지능 혁신에 따라 더욱 심화될 허위정보(disinformation)의 위협을 분석함. 특히 허위정보를 통한 영향공작과 선거개입이 생성 인공지능을 통하여 어떻게 진화할 것인지를 과거 및 현재 사례를 중심으로 살펴 봄. 이 문제에 대한 대응 방향으로 ① 대응 주체 차원에서의 노력 ② 기술적 대응 방안 ③ 제도적 대응 방안 ④ 글로벌 공조 등을 제시하였음

Interference on election by AI in the age of disinformation

Summary

This report investigates the new threat of disinformation in the age of generative AI. In particular, the possible evolution path of influence operation and interference on election is discussed via past and present episodes. The following countermeasures are suggested : ① the efforts of each actors in the AI ecosystem, ② technical countermeasures ③ regulatory countermeasures ④ global cooperation

1. 정보공간에서의 영향 공작(Influence Operation)과 선거개입

◆ 정보공간을 둘러싼 경쟁의 심화와 영향공작(Influence Operation), 선거 개입

- 온라인 공간, 특히 소셜 미디어는 이미 강대국간 담론 경쟁이나 상대국의 사회적 신뢰를 약화시키는 공간으로 활용되어 왔음
 - 특히 정보의 자유로운 흐름 또는 통제에 인공지능의 역할이 커지면서 정보·데이터·콘텐츠를 좌우하는 인공지능의 악용에 대한 우려가 확산
 - 세계경제포럼의 Global Risks Report 2024(2024. 1. 10. WEF)는 인공지능의 위험, 특히 오정보 및 허위정보(disinformation)를 새로이 대두하는 글로벌 리스크 가운데 하나로 강조¹⁾
- 국가 차원의 허위정보를 영향공작, 특히 선거개입의 수단으로 활용하는 것이 기술 혁신으로 점차 용이해지고 있음
 - 영향공작에는 자국에 유리한 담론 확산, 온·오프라인 선전, 인지전(cognitive warfare), 선거 개입, 해외 우호 그룹의 형성과 이용이 모두 포함됨
 - 틱톡을 둘러싼 논란도 개인정보 보호 차원을 넘어서서, 틱톡의 미디어로서의 영향력을 차단하려는 노력을 일환으로 이해할 수 있음

1) WEF, Global Risks Report 2024 (2024. 1. 10)

- 생성형 인공지능 혁신은 기존 온라인 공작의 저비용 자동화라는 측면에서 영향 공작, 프로파간다에 큰 변화를 가져올 수 있음

◆ 주요 사례

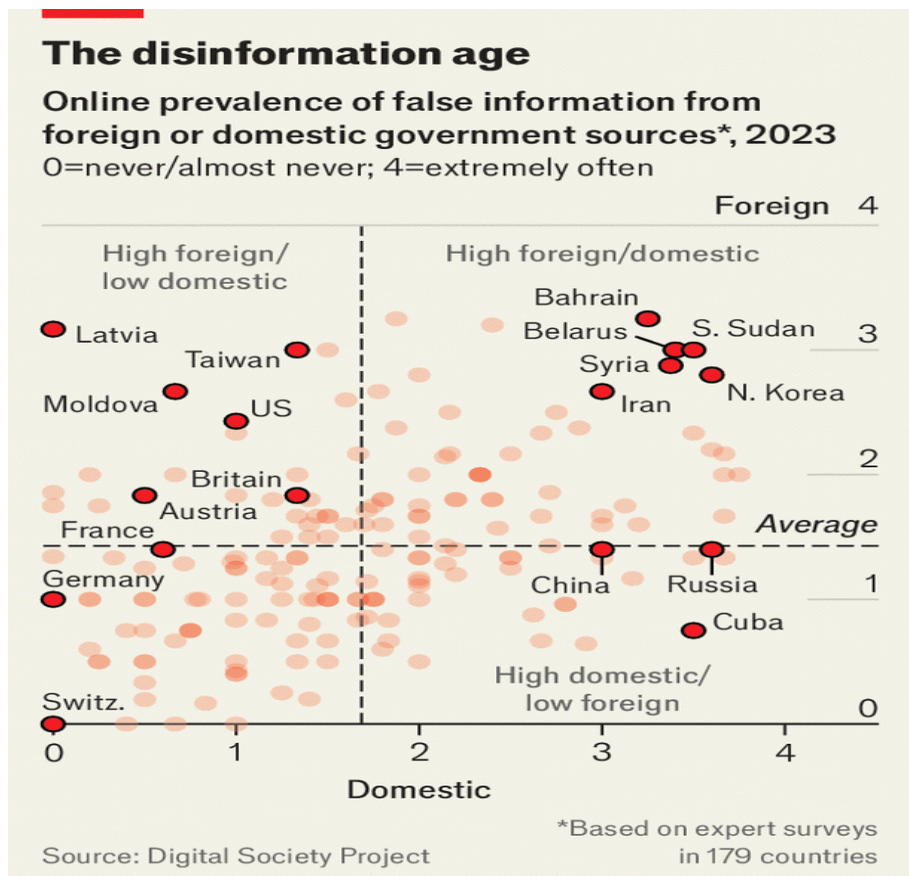
- 서방의 정보기관 및 미디어는 러시아와 중국을 국가 단위에서의 영향공작 및 선거 개입을 수행하는 대표적인 국가들로 지목하고 있음
 - 러시아와 중국은 기술기업에 대한 통제력, 대외적인 우호세력 확보 및 담론 경쟁의 우위에 대한 우선순위 부여라는 특징을 보유²⁾
 - 주로 소셜 미디어의 허위계정을 통하여 의도적으로 잘못된 정보를 유포함으로써 장기적·간접적으로는 담론 경쟁에서의 우위, 단기적·직접적으로는 선거에 영향을 미치고자 함
 - 최근의 전형적인 허위정보 유포 방식은 자신들이 확보한 계정에 허위 정보를 심어두고(Plant), Medium, Reddit 등 수 많은 팔로워를 가진 계정으로 증폭시키고(layering), 신뢰할 수 있는 사이트가 이를 인용하면서(integration) 완성됨
 - 많은 계정들이 저소득 국가에서 생성, 판매되고 있음
 - 허위계정 간의 조율된 행위(co-ordinated inauthentic behaviour: CIB)는 큐레이션 알고리즘이 허위정보를 인기 있는 것으로 인식하게 해

2) 특히 중국은 특유의 통일전선 기술을 통하여 영향공작에 유리한 정치·사회적 환경 조성을 추구해왔다고 평가됨. 통일전선 기술은 국내의 주요 사회 세력과 유력인사를 통합해 우호 세력을 확보하고 적을 내부에서부터 약화시키는 기술로 구체적으로는 특정국 미디어 장악, 해외거주 중국인 디아스포라 동원 체제, 온오프라인에서의 여론, 선전 활동, 허위정보 유포 등을 통하여 우호세력을 구축, 일단 통일전선이 구축 되면 편향적 담론의 확산뿐만 아니라 해당국의 사회적 분열과 갈등 조장, 정치적 영향력 행사, 미디어 혼탁, 기술 탈취와 같은 산업스파이 활동을 망라해 상대방에 대한 우위를 추구

악의적 정보를 확산시키는데, 정부와 직접 연관이 없는 트롤 공장, 마케팅 회사 등을 이용하므로 상기의 매커니즘을 발견, 조치하기에 어려움

- 주요 타겟인 국가로는 미국과 대만이 대표적인
- 글로벌 차원에서의 담론 경쟁, 통일전선 차원에서의 영향력 강화가 주된 이유일 것으로 판단됨

<도표 1> 정부 차원에서 유포된 허위정보 대상국 현황



출처 : Economist, Disinformation is on the rise, How does it work? (2024, 5, 1)

- Project Lakhta라 지칭되는 미국 대선 개입이 러시아에 의한 선거 개입의 대표적인 사례
 - 미국의 2016년 대선에서 상트페테르부르크 소재 인터넷 리처치 에이전시 (IRA)가 수천 개의 소셜 미디어 계정을 만들어 親트럼프, 반클린턴 허위정보, 가짜 뉴스를 외부로 확산
 - 스텐포드 대학 사이버 정책 센터의 보고서는 러시아의 선거 개입 사례분석을 통하여 미래 선거가 실제 득표수 조작에서부터 딥페이크나 인공지능 콘텐츠 생성 기술의 활용에 이르기까지 다양한 문제들에 직면할 것이라고 경고³⁾

- 중국의 영향 공작과 선거 개입
 - 주요 수단은 허위 계정을 이용한 선전 활동과 정보 유통으로, 드래곤브릿지(DragonBridge), Storm-1376 또는 스팸모플라지 (Spamouflage)로 알려진, 여러 플랫폼에 걸친 수많은 허위 계정 네트워크를 이용하는 것으로 알려져 있음⁴⁾
 - 알려진 사례 가운데 영어, 독일어 등 다국적 언어를 동원한 정보의 유포도 자주 발견되며, 여기에는 한국어도 포함
 - 호주전략정책연구소에 따르면 2017년 8월 ~ 2022년 12월 기간 중 외부 세계에 대한 온라인 영향공작으로 적어도 51건의 보고 사례가 알려져 있는데 이용 플랫폼은 트위터, 페이스북, 레딧, 유튜브,

3) 'Securing American Elections: Prescriptions for Enhancing the Integrity and Independence of the 2020 U.S. Presidential Election and Beyond' 참조할 것. Stanford Cyber Policy Center, (2019. 6. 6)

4) 'Pro-PRC Influence Campaign Expands to Dozens of Social Media Platforms, Websites, and Forums in at Least Seven Languages, Attempted to Physically Mobilize Protesters in the U.S.' (2021. 9. 7) 참조할 것

인스타그램 등을 망라하였다고 함⁵⁾

- 중국의 선거 개입 의혹은 미국, 대만, 태국, 캐나다, 호주 등 많은 국가에서 제기되었는데, 특히 대만에 대한 영향공작 및 선거개입에 대한 지적이 다수
- 조지 워싱턴 대학 아시아학 연구소 보고서(2019.1)에 따르면 중국의 댓글부대인 50센트당이 대만의 주요 소셜 미디어에서 하루 평균 2500회 반중(反中)후보를 공격⁶⁾
- 2023년 10월 차이밍옌(蔡明彥) 대만 국가안전국 국장은 의회 보고에서 중국에서 만든 가짜 뉴스를 대만 국가안전국이 적발해 정부에 통보한 사례가 1700건에 달하며 총통선거에서 여론조사 기관이나 광고회사를 이용해 중국이 여론조사를 조작할 가능성을 경고⁷⁾

5) ASIP, "The Party Speaks for You", (2020. Policy brief No. 32) 및 A. Zhang, 'Gaming Public Opinion', ASPI, (2023. 리포트 71호) 참조할 것

6) Sigur Center for Asian Studies, 'Countering China's Sharp Power: Disinformation and Social Media in Taiwan', Asia Report No. 44 (2019.1)

7) 중앙일보, '대만 "중국 정부, 선거에 개입할 방법 많아"...여론조작 경고 나섰다' (2023. 10. 4)

2. 인공지능 혁신과 위협

- 생성형 인공지능 등 인공지능 혁신은 허위정보 생성 및 유포에 있어서의 저비용과 확장성으로 인해 선거 개입 등 정보공간에서의 영향 공작을 새로운 차원으로 발전시킬 수 있다는 중대한 위협을 제기
 - 생성형 인공지능은 현재 LLM 모델이 앞서나가고 있지만 오디오, 비디오, 음악 콘텐츠도 가능해지면서 진실과 거짓 간의 경계는 더욱 불분명해질 것임
 - 이미 최초의 국가 단위 인공지능 생성 콘텐츠를 이용한 해외선거 개입사례로 MS의 Threat Intelligence Team이 2024년 1월 대만 총통 선거를 지목⁸⁾
 - 스텐포드 인터넷 관측소, 조지타운대학 안보센터 및 OpenAI는 생성형 인공지능이 영향 공작에 미치는 영향을 모델 참여자, 행위, 콘텐츠 자체로 구분하여 아래 표와 같이 제시

〈표 1〉 생성형 인공지능과 영향 공작

	생성 AI로 인한 잠재적 변화	변화 내용
AI 참여자	보다 큰 규모의, 다양한 프로파간다 행위자 대두	프로파간다 생성 비용의 감소로 더많은 주체가 프로파간다 활동에 참여
	아웃소싱 기업이 더욱 중요해짐	프로파간다 콘텐츠를 자동화하는 고용인력의 경쟁력 강화

8) Economist, Disinformation is on the rise. How does it work? (2024. 5. 1) 참조

	생성 시로 인한 잠재적 변화	변화 내용
행 위 (Behavior)	자동 콘텐츠 생성으로 선전활동의 규모 증대	자동화로 인한 비용 절감으로 프로파간다 활동이 보다 용이해짐
	현재의 선전 행위가 보다 효율적으로 발전	다양한 플랫폼을 아우르는 테스트 (cross-platform test)와 같은 고비용 전술이 저비용으로도 가능
	새로운 전술 등장	동태적, 실시간, 개인화된 콘텐츠 생성 등
콘텐츠	선전 메시지가 더 신뢰할 수 있고 설득력 있게 진화	과거에 비해 더욱 전문적이고 믿음직한 메시징이 가능
	프로파간다를 발견하기 더욱 어려워짐	단순 복사/붙임 차원을 넘어선, 보다 정교한 메시징 등

출처 : J. Gokdstein 등, 'Generative Language Models and Automated Influence Operation(조지타운 안보센터, 스탠포드 인터넷 관측소, OpenAI, 2023, 1)

- 인공지능이 개인정보를 활용하여 개인이나 소집단애의 실시간 마이크로 타게팅(Micro Targeting)을 수행할 수도 있음
- 특히 특정 사안에 풀뿌리 운동처럼 보이려고 공적 관계나 정치적 캠페인을 이용하는 에스트로터핑(astroturfing)이 자동화된 프로그램으로 가능할 수도 있음
 - 랜드(RAND) 연구소의 보고서는 메시징 자체보다 메시지가 그럴듯하게 보이는 능력을 갖춘 생성형 인공지능의 강점에 주목하면서 에스트로터핑의 위협을 강조하고 이를 소셜 미디어 조작 3.0 시대로 지칭⁹⁾
 - 에스트로터핑 혁신으로 특정 정부나 집단이 자체 미디어를 저비용으로

9) RAND, The Rise of Generative AI and the Coming Era of Social Media Manipulation 3.0 (2023, 9)

운용하면서 평상시부터 타겟이 되는 국내외 소셜 미디어 이용자와 소통하며, 학습을 통해 모델의 정교화가 가능

- 유비쿼터스한, 거대한 봇 네트워크가 텍스트, 이미지, 비디오, 오디오 콘텐츠를 생성해 소셜 미디어에서 메신저의 신빙성(authenticity)을 지원

〈표 2〉 소셜 미디어 조작의 세대별 구분

세대	핵심 기술	예시
1.0	기초적 프로그래밍	반자동 봇이 인간이 생성한 콘텐츠를 포스팅 (50센트 당)
2.0	초기 기계 학습	저사양 조작 비디오, 제한적 컴퓨터 생성 콘텐츠, 봇을 통한 배분
3.0	생성형 인공지능	고사양의 tailored 텍스트, 이미지 at scale, 동태적, 자동화된 배분 및 coordination via Bots (for astroturfing)

출처 : RAND, The Rise of Generative AI and the Coming Era of Social Media Manipulation 3.0 (2023. 9)

- 랜드 연구소의 보고서는 중국의 인민해방군 영향공작 이론가인 리비청(李弼程)의 이론도 소개
 - 이에 따르면 공작지휘 조직으로부터의 지도와 특정 캐릭터를 갖춘 가공의 페르소나, 온라인 여론 상황에 따른 조정, 다수의 페르소나 간의 조율·협력이 작동해 소셜 미디어 여론을 이끄는 인공지능 모델을 통해 미래 정보전, 인지전(Cognitive Warfare)을 중국이 수행해야 함을 설파
 - 만약 이런 생성 인공지능 모델이 작동한다면 생성 인공지능의 에스트로터핑은 평소에 잘 드러나지 않다가 결정적 순간에 정치적 나레이션이 가능할 것이고 선거에도 큰 영향을 미칠 수 있을 것임

3. 대응 방향과 시사점

- 인공지능을 통한 영향 조작, 선거 개입은 모든 국가에 잠재적인 위협이 될 수 있다는 경각심이 요구됨
 - 주요 사례로, 중국 언론홍보업체 하이마이와 하이준이 서울프레스·부산온라인·전라오늘 등 국내 언론 위장 사이트 38개를 통하여 친중·반미 콘텐츠를 무단 유포했다는 사실이 최근 국내 미디어에 대대적으로 보도됨¹⁰⁾
- ‘중국 정부의 코로나 공조 성과’, ‘한국의 민주주의 정상회의 참석, 득보다 실이 많다’ 등 무단 유포된 콘텐츠는 진위의 판별이 용이한 가짜 뉴스가 아니라 일종의 의견으로, 중국에 이익이 되는 여론 형성에 기여할 수 있는 이야기 전쟁 또는 담론 경쟁의 일환

◆ 대응 주체 차원에서의 노력

- 연구기관, 플랫폼 기업, 정부, NGO, 미디어 기업간의 협조가 중요
 - 특정 이용자나 포스트를 차단 또는 라벨링, 미디어 리터러시 교육, 플랫폼들에의 입증 책임(verification requirement) 강화, 사실 확인(fact-checking) 기관이나 조직 지원 등
 - 각 주체 간의 정보 및 인력 공유, 특히 허위정보 공격에 대한 데이터 접근이 중요¹¹⁾

10) 조선일보, ‘국정원 “中 업체, 국내 언론 위장 사이트 38개 개설... 친중반미 콘텐츠 유포” (2023. 11. 13)

- 생성형 인공지능 위험은 모델 구축, 접근, 콘텐츠 유포 및 신뢰 형성 과정의 단계로 구분해 각 주제가 대응하는 것이 합리적

- OpenAI는 단계별로 정부, 기업, 개발자, 이용자 등 인공지능 참여자별 주요 단계별 대응 방향을 아래 표와 같이 제시¹²⁾

〈표 3〉 OpenAI가 제시한 생성형 인공지능 서비스 제공 이용 단계별 대응방안

단 계 (Stage)	1. 모델 구축	2. 모델 접근 (access)	3. 콘텐츠 유포	4. 신뢰 형성 (Belief Formation)
위험 경감 (Mitigation) 조치	사실(fact)에 입각하는 AI 모델 구축	AI 제공자가 언어 모델에 강한 이용 제한을 부여	플랫폼 및 AI 제공자가 AI 콘텐츠를 식별에 협력 (coordinate)	주요 기관의 미디어 독해 (literacy) 캠페인 참여
	생성 모델의 추적을 위해 개발자가 데이터를 마크하고 입증할 수 있는 데이터를 확산	AI 제공자가 모델 출시의 새로운 규범을 발전시킬 것	플랫폼이 인격증명 (proof of personhood)를 요구	이용자를 위한 AI 도구(tool)를 개발자가 제공
	데이터 수집에 대한 정부 규제	AI 제공자의 안전 취약점 봉쇄	공공 데이터 이용자는 잘못된 콘텐츠에의 노출을 축소	
	AI 하드웨어에 대한 접근을 정부가 통제(control)			

출처 : Forecasting potential misuse of language models for disinformation campaigns and how to reduce risk, (OpenAI, 2023. 1. 11)

11) EU의 디지털 서비스 법안(DSA)은 플랫폼들이 사회에 대한 구조적 리스크(systemic risk)에 대응토록 연구기관에 데이터를 제공하도록 함

12) Forecasting potential misuse of language models for disinformation campaigns and how to reduce risk, (OpenAI, 2023. 1. 11)

◆ 기술적 대응 방안

- 인공지능은 허위정보 생성, 유포뿐만 아니라 방어 수단으로도 기능할 수 있어 관련 기술개발이 필요
 - 빅테크들은 허위 계정간 관계 분석 등 패턴 분석으로 공격자 및 콘텐츠를 식별(Imposter-detection 알고리즘)
 - TrueMedia.org와 같은 민간 사이트들이 추적 도구도 제공
- 신뢰할 수 있는 인공지능 시스템 및 리스크 관리 관련 표준, 기술 및 인력 양성 지원
 - 특히 모니터링 및 측정 관련 방법, 기술의 개발 지원
 - 가짜뉴스, 허위정보를 막는 인공지능 기술 및 인력양성 지원 강화

◆ 제도적 대응 방안

- EU와 호주는 이미 허위정보 관련 규제를 본격적으로 시행하고 있음
- EU의 허위정보에 대한 실행 규약 (Code of Practice on Disinformation) 주요 내용
 - 허위정보를 통한 금전적 이익을 도모하는 웹사이트 차단 또는 유예
 - 정치광고나 특정 사회적 이슈에 기반하는 광고의 라벨링
 - 허위계정 탐지를 위한 기술이나 정보 출처의 신뢰성 등급부여 기술 등을 장려

- 연구자나 사실확인 커뮤니티(fact-checker)의 플랫폼 기업 데이터 접근 및 미디어 독해력(literacy) 운동 지원 등의 실행 규약을 시행
- 참여 기업들은 연간 자체 평가 리포트를 제공하고 EC가 점검하도록 함
- 호주의 허위정보 및 오정보에 대한 실행 규약 (Australian Code of Practice on Disinformation & Misinformation)
 - 디지털산업그룹(DIDG) 주도의 민간 자율 규약으로, 허위정보, 오정보의 확산에 대처하기 위한, 참여 디지털 플랫폼의 최소한의 약속 내지는 책무(minimum commitments)를 명시
 - 주요 조치 내용
- 이용자 행위 및 콘텐츠에 대한 인간에 의한 검토(human review) 정책/절차를 수행
- 잘못된 정보에 라벨링(labeling)을 하거나 정보의 신뢰 수준 표시(trust indicator)
- 허위/오정보의 검색 순위 강등(demoting)
- 진실되지 않은 행위에 의해 유포된 콘텐츠 삭제
- 해당 행위 이용자 계정의 유예 또는 폐쇄
- 콘텐츠의 허위여부, 행위의 진실성 여부, 출처의 진실 여부 등을 식별할 수 있는 기술적 해법 제공
- 명확한 편집 정책 및 콘텐츠 기준을 공표

- 현재 트위터, 페이스북, 구글 등 거대 플랫폼들이 참여하고 있으며 주요 플랫폼 기업들은 자신들의 콘텐츠 모더레이션(moderation) 정책에 기반해 자율성을 일정 수준 보유¹³⁾
- 우리도 민간 자율 규약에 대한 논의를 검토할 필요
 - 직접적인 규제보다는 민간의 콘텐츠 모더레이션 정책 보완이나 기업간 협력 사안 등을 검토

◆ 글로벌 공조

- 인공지능은 특정국내에서의 거버넌스로는 다양한 위험, 영향을 통제할 수 없어, 글로벌 공조가 불가피
 - 글로벌 플랫폼 기업들이 글로벌 이용자들을 대상으로 서비스를 제공하고 있음
- 특정국이나 집단의 영향공작, 허위정보를 통한 간여도 글로벌 차원에서 이루어질 수 있음
 - 주요 사례로 미국-EU 무역기술위원회, 히로시마 AI 프로세스 등이 있음
- 미국-EU 무역기술위원회 로드맵 (TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management, 2022. 12)

13) 단, 규약 의무에서 일부 면제되거나 독자적인 조치를 취할 경우 이에 대한 합리적 설명이 제시되어야 함

- 공통의 용어(terminology) 및 분류, 국제 표준, 신뢰할 수 있는 AI 및 위험관리를 위한 도구(tool)의 분석 및 수집, 현재 및 미래 리스크의 모니터링·측정·추적 등 세 가지 영역을 중심으로 전문가 실무그룹을 구성해 협력 추진
- 히로시마 인공지능 프로세스
 - G7 실무그룹이 생성 인공지능 및 인공지능 정책 전반에 관한 글로벌 공조방안을 제시하기로 합의
 - 공조·협력의 주요 의제에는 일반적인 거버넌스 정책뿐만 아니라 인공지능 관련 IP의 보호, 컴퓨팅 기술 분야 R&D 협력, 허위정보 등 외국의 정보 조작(manipulation)에의 대응을 명시
- 캠프데이비드 한·미·일 정상들의 공동성명(2023. 8. 19)
 - 인공지능이 안보와 관련하여 제기하는 도전과제에 함께 대응하기로 선언
 - 특히 해외 정보 조작과 감시 기술의 오용, 허위 정보의 위협에 대한 대응능력 조율을 강조

참 고 문 헌

[국내문헌]

- 조용래, '新안보시대 국가혁신정책의 새로운 엔진, 국가전략기술', 과총웹진 (2023. 6. 5)
- 외교부, G7 정상 공동성명 (2023. 8. 18)
- 조선일보, '국정원 “中 업체, 국내 언론 위장 사이트 38개 개설... 친중·반미 콘텐츠 유포” (2023. 11. 13)
- 중앙일보, '대만 “중국 정부, 선거에 개입할 방법 많아”...여론조작 경고 나셨다' (2023. 10. 4)

[해외 문헌]

- A. Zhang, 'Gaming Public Opinion', ASPI, (2023. 리포트 71호)
- J. Gokdstein 외 'Generative Language Models and Automated Influence Operation', 조지타운 안보센터, 스텐포드 인터넷 관측소, OpenAI, (2023. 1)
- RYAN SERABIAN, LEE FOSTER, 'Pro-PRC Influence Campaign Expands to Dozens of Social Media Platforms, Websites, and Forums in at Least Seven Languages, Attempted to Physically Mobilize Protesters in the U.S.' (2021. 9. 7)
- ASIP, "The Party Speaks for You", (2020. Policy brief No. 32)
- DIDG, 'Austalian Code of Practice on Disinformation & Misinformation', (2021. 2)
- Economist, Disinformation is on the rise. How does it work? (2024. 5. 1)
- EU, 허위정보에 대한 실행 규약 (Code of Practice on Disinformation) (2018. 4)
- OpenAI, Forecasting potential misuse of language models for disinformation campaigns and how to reduce risk, (2023. 1. 11)

- RAND, The Rise of Generative AI and the Coming Era of Social Media Manipulation 3.0 (2023. 9)
- Sigur Center for Asian Studies, 'Countering China's Sharp Power: Disinformation and Social Media in Taiwan', Asia Report No. 44 (2019.1)
- Stanford Cyber Policy Center, 'Securing American Elections: Prescriptions for Enhancing the Integrity and Independence of the 2020 U.S. Presidential Election and Beyond' (2019. 6. 6)
- TTC, 'TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management', (2022. 12)
- WEF, Global Risks Report 2024 (2024. 1. 10)