

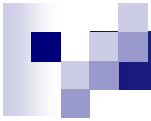
ディープフェイクに関する 各国の規制の動向

湯浅 壱道

明治大学公共政策大学院ガバナンス研究科教授



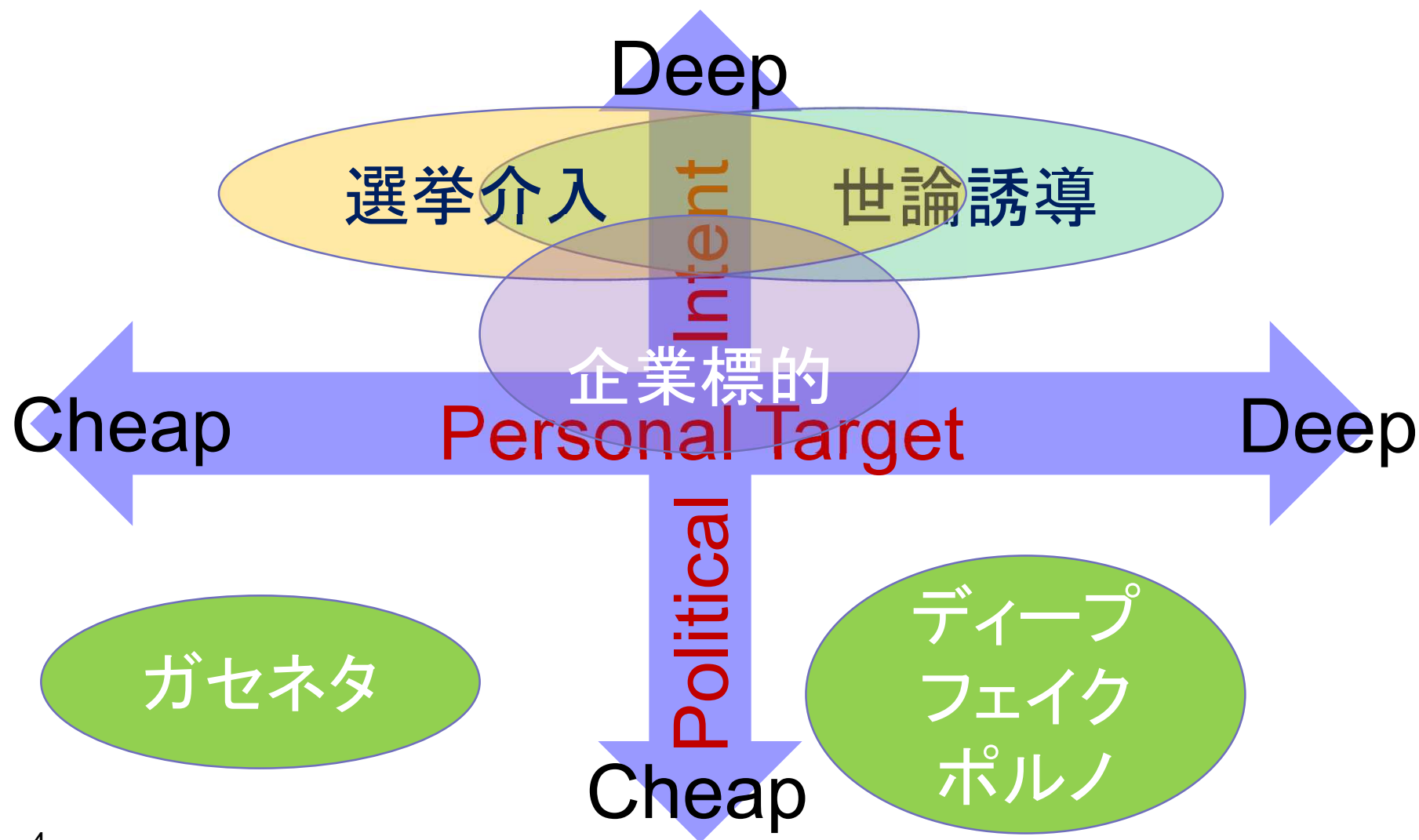
現状

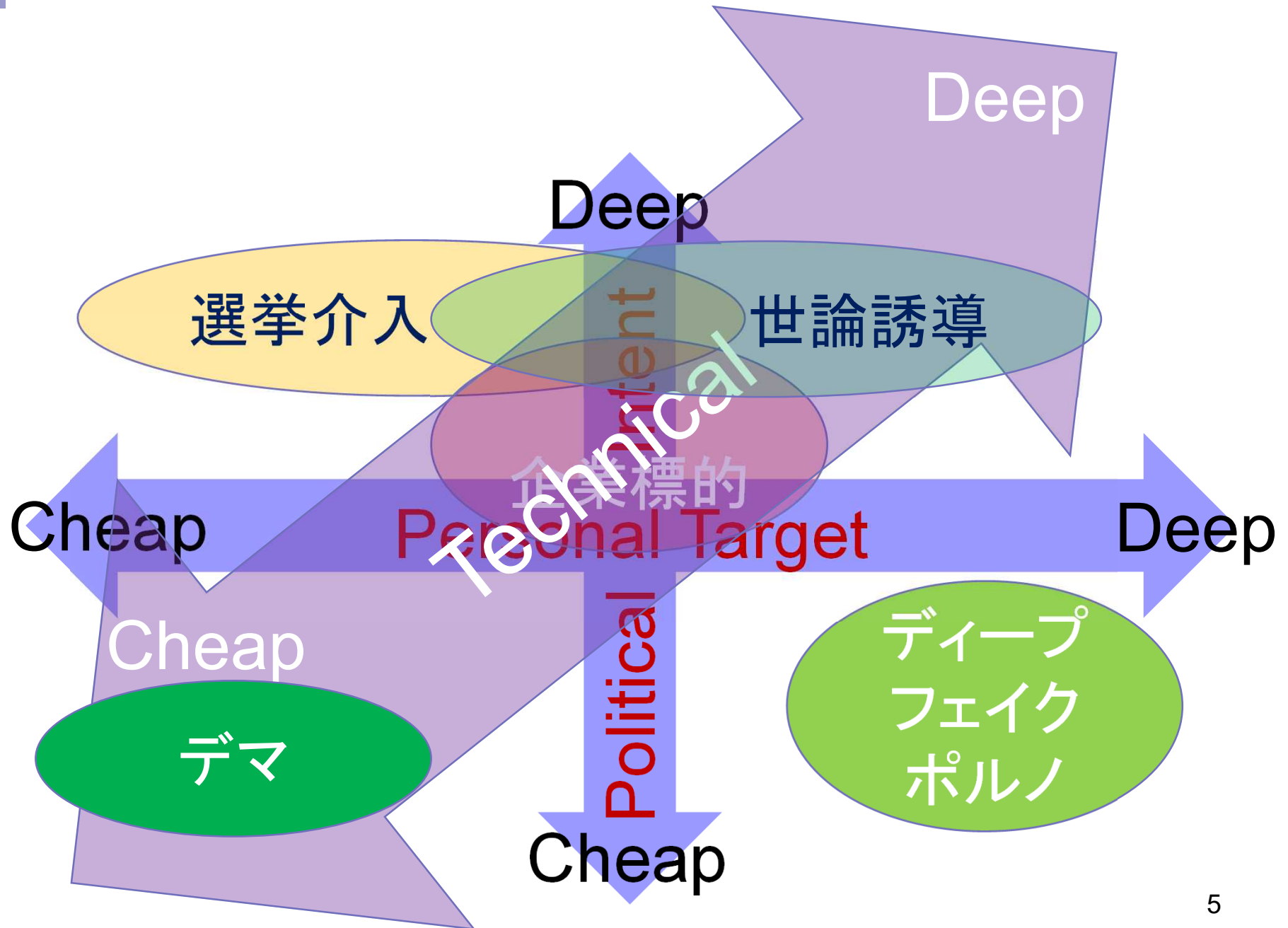


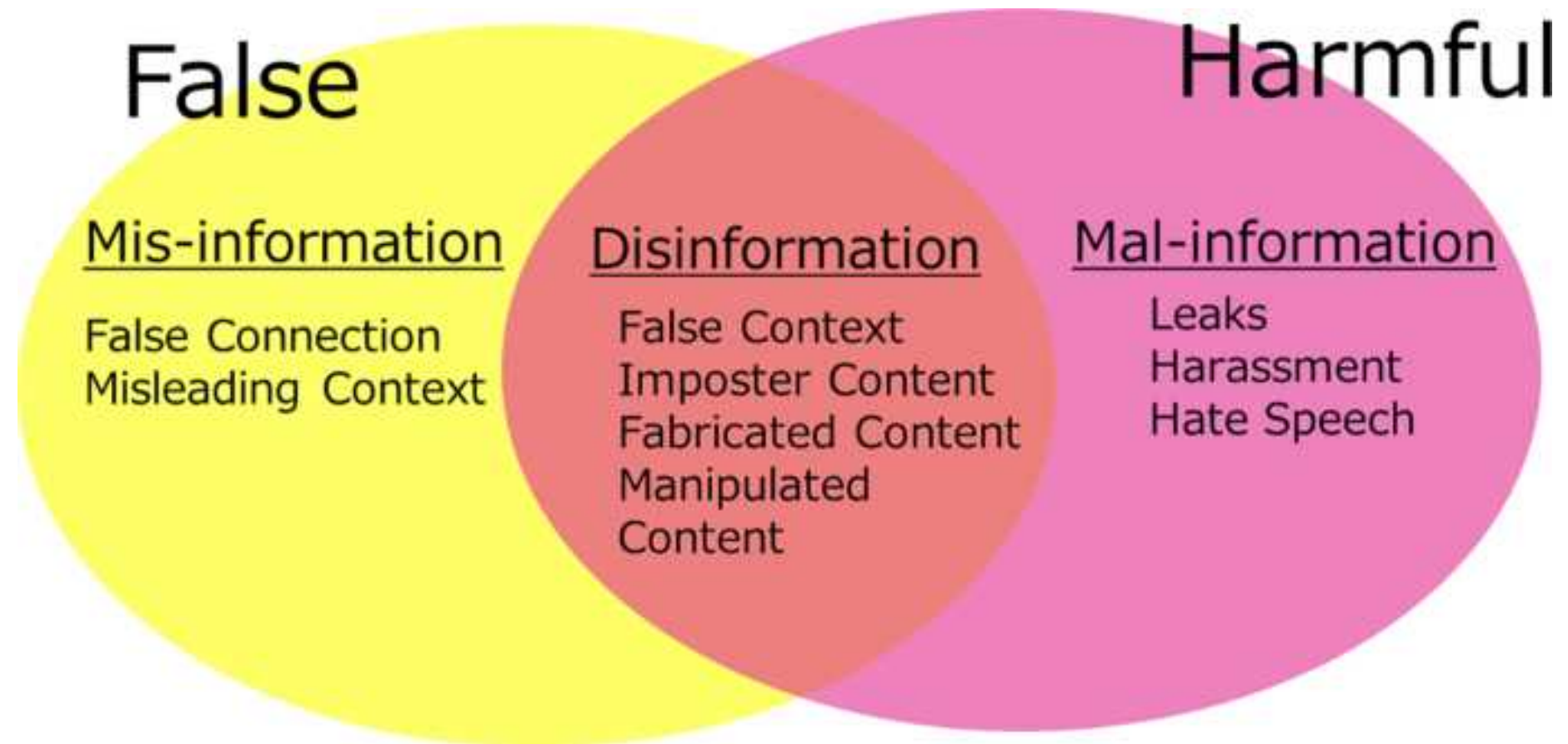
SNSと民意形成

- フィルター・バブル
- フェイクニュース・ディスインフォメーションの
流布とその規制
- プラットフォーマーによるアカウント凍結や表
現抑止の是非
- 特定の個人を標的として民意誘導を狙うマイ
クロターゲティング
- インターネット広告(アドテクノロジー)の高度
化

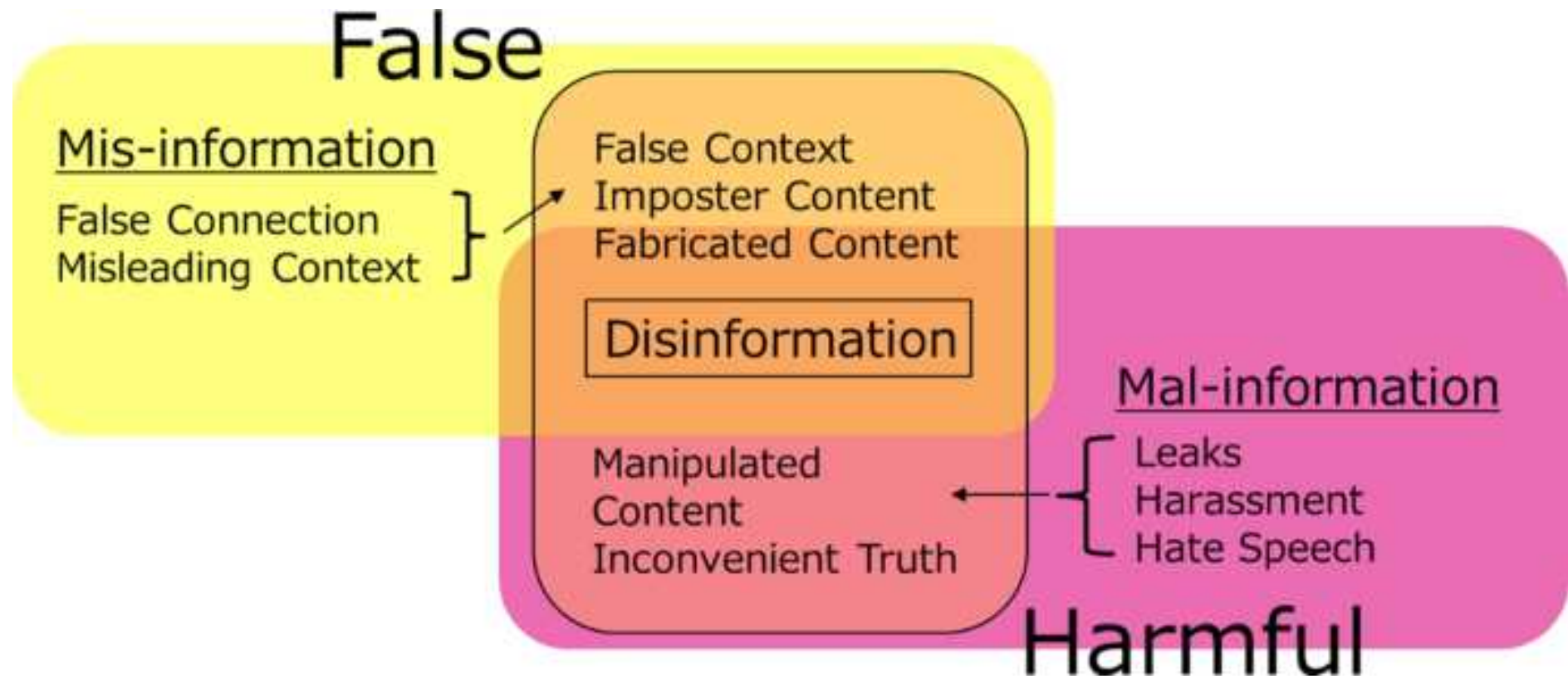
Fakenews, Disinformation







Source: Nagasako 2020



Source: Nagasako 2020



▲...Russia (31 cases) ●...China (6 cases) ◆...Unknown

Source: Nagasako 2020



サードパーティーCookie禁止の動き

■ Google Chrome

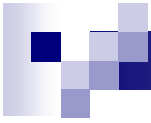
- Chrome におけるサードパーティ Cookie のサポートを段階的に廃止
- <https://privacysandbox.com/>

An evolution of the web

Privacy is at the forefront of public dialogue as people seek to understand and control how their personal information is used online.

Some of the information collected by sites and third-parties is necessary to provide the rich content and services you expect. However, the tools used to provide this have gone far beyond their original intent in their ability to recognize you, your online activity and the devices you use.

The Privacy Sandbox will lay new foundations for a safer, more sustainable, and more private web.



- Privacy is at the forefront of public dialogue as people seek to understand and control how their personal information is used online.
- Some of the information collected by sites and third-parties is necessary to provide the rich content and services you expect. However, the tools used to provide this have gone far beyond their original intent in their ability to recognize you, your online activity and the devices you use.
- The Privacy Sandbox will lay new foundations for a safer, more sustainable, and more private web.



Prevent tracking as you browse the web.

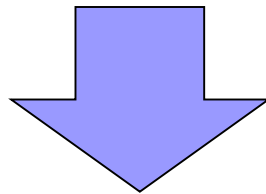
People should be able to browse the web without worrying about what personal information is being collected, and by whom.

The Privacy Sandbox initiative aims to remove commonly used tracking mechanisms, like [third-party cookies](#), and block covert techniques, such as [fingerprinting](#).

<https://privacysandbox.com/>

フィンガープリント

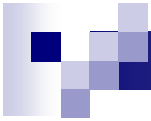
- サードパーティーCookieに代わる技術
- ユーザーが使うブラウザやデバイス(端末)の動作環境の特徴を、フィンガープリント(指紋)のように手掛かりとする



- ネット上の行動を追跡
- 9割以上の精度で特定(?)



規制の態様



対策

■ EU

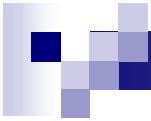
- ディスインフォメーション対策全般、ディープフェイクには特化せず
- プラットフォーマーに対する自主的規制(Code of Practice)要求、共同規制(コーポラティズム的)

■ アメリカ州法(選挙法)

- 選挙におけるディープフェイク規制

■ アメリカ連邦法

- 安全保障の観点からディープフェイク研究



※ディープフェイクと国際法

■ タリン・マニュアル

- 「武力行使」該当性と自衛権行使の要件

■ サイバー空間を利用した選挙干渉は？

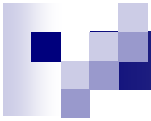
- 政治的言説として「戦争」

- ディック・チェイニー元副大統領、ニッキー・ハーリー国連大使 等

- 諜報活動の一環、法的には戦時国際法は適用されないという見方が通説



アメリカにおける規制



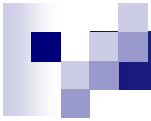
重要インフラ指定

- 2017年1月6日
- 選挙管理システムが国土安全保障省により重要インフラストラクチャー指定
 - 有権者登録データベース及び関連する情報通信システム
 - 選挙管理に使用される情報通信インフラ及びシステム（投票結果の開票、集計及び表示システム、選挙後の選挙結果検証報告用のシステムなど）
 - 投票システム及び関連するインフラ
 - 選挙管理及び投票システム用のストレージ装置
 - 期日前投票所を含む投票所

- 国土安全保障省

- 州の要請に応じて選挙システムのセキュリティに関する支援を行う
- 選挙システムに関し、情報共有及び分析センター(ISAC)として選挙インフラISAC(Elections Infrastructure ISAC = EI ISAC)設置
- 連邦政府と州(地方政府)が情報共有、分析、連携

- 政治活動委員会(PAC)、選挙運動自体、政府や州政府等が設立したものではない選挙関係団体は、重要インフラストラクチャーに含まれない
- 政党、候補者等については、重要インフラ指定外



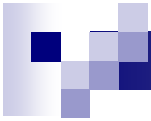
カリフォルニア州法

- 2019年に選挙運動におけるディープフェイク等を規制するAB730法案を可決
- 投票日の60日前から、何人にも候補者に関するディープフェイクの発信を禁止
- 2023年1月1日までのサンセット法
- 規制対象
 - 「実質的に虚偽の音声または視覚メディア (materially deceptive audio or visual media)」
画像、音声または動画が対象



■ フェイクの定義

- 「候補者の外観、スピーチ、または行為に関する画像、音声または動画であって、当該の画像、音声または動画が誤って表示されるような方法で意図的に操作されたもの」、または「合理的な人物が、変更されていない元のバージョンを聞いたり見たりした場合に、その人物が持っているものよりも、画像、音声または動画の表現内容について根本的に異なる理²⁰解または印象を持つようにしたものの」



20010.

(a)

(b)項に規定されている場合を除き、いかなる人、協会、団体、企業、選挙運動委員会または選挙組織も、現実の悪意を持って、次の選挙運動用資料を製造、配布、公開または放送してはならない。

(1)公職の立候補者の画像または写真が重ねられているもの

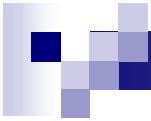
(2)公職の立候補者の画像または写真であって、他の人の画像または写真が重ねられているもの

本条でいう選挙運動用資料とは、印刷物、新聞広告、その他の定期刊行物、テレビコマーシャル、コンピューター画像などが含まれるが、これらに限定されない。本条でいう現実の悪意とは、人の画像が写真または写真に重ねられて虚偽の表現を作成したという知識、または人の画像が重ねられたかどうかを虚偽の表現のために無謀に無²¹



(b)

個人、協会、団体、企業、選挙運動委員会または組織は、選挙運動用資料で用いられている文字と同等のポイント大で「この画像は事実を正確に表現するものではありません」という文言を表示する場合に限り、(a)項で禁止されている画像または写真を含む選挙運動用資料を作成、配布、公開、または放送することができる。当該の表示は、(a)項で禁止されている各画像または写真のすぐ隣に配置しなければならない。



テキサス州法

- 2019年に選挙法を改正し、選挙運動におけるディープフェイク等を規制するSB751法案を可決
- ディープフェイクという文言を用いると共に、候補者を誹謗中傷したり選挙結果に影響を与えたりすることを目的としてディープフェイク・ビデオを作成することが明確に禁止
- 投票日の前30日間はフェイクと知りつつ流布させることも禁止



255. 004条 真実の情報源

(a) 候補者を傷つけたり、選挙結果に影響を与えたりする目的で、真実の情報源以外の源からから発せられた政治広告を印刷、公開、または放送する契約またはその他の契約を締結した場合、その人物は犯罪を為すものとする。

(b) 候補者を傷つけたり選挙の結果に影響を与えたりする意図で、選挙運動の通信において、当該通信がその真の出所以外の出所から発せられていることを知りつつ故意に表現した場合、その人物は犯罪を為すものとする。

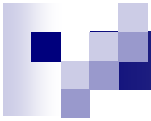
(c) この節の違反は、A級の軽犯罪とする。

(d) 候補者を傷つけたり、選挙結果に影響を与えたりすることを意図して、次の行為を行った場合、その人物は犯罪を為すものとする。

(1) ディープフェイク・ビデオを作成すること

(2) 選挙の前30日以内にディープフェイク・ビデオを公開または配布すること

(e) 本節でいう「ディープフェイク・ビデオ」とは、欺くつもりで作成され、実際には発生しなかった行動を実行している実際の人物を描写しているかのように見えるビデオをさすものとする。



連邦法

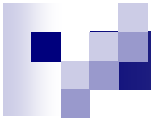
- 2020年連邦議会で可決

- 敵対的生成ネットワークの出力の識別に関する法律 (Identifying Outputs of Generative Adversarial Networks = IOGAN法)

- 「敵対的生成ネットワーク」(Generative Adversarial Networks = GAN)

- 2021年国防授權法

- 国防長官は「軍人とその家族を対象としたサイバー搾取とオンライン詐欺に関する研究」を行うことが義務付けられた



IOGAN法

- 「全米科学財団理事長に、敵対的生成ネットワーク(ディープフェイクとも呼ばれる)、将来開発される可能性のあるその他の同等の技術、およびその他の目的でに成される可能性のある出力に係る研究に対する支援を行うことを命ずる法律」
 - 全米科学財団(NSF)および米国国立標準技術研究所(NIST)に対し、敵対的生成ネットワークによって出力されたものを含む操作または合成されたメディアに関する研究を支援するよう指示 26

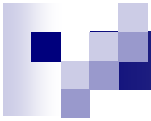


■ NSF

- 操作・合成されたコンテンツや情報の真正性に関する研究を支援する

■ NIST

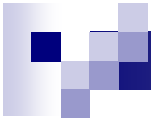
- 敵対的生成ネットワークやコンテンツを合成・操作する他の技術の機能や出力を調べるための技術ツールの開発を加速するため、必要な測定法や基準の開発のための研究を支援しなければならない



■ NSFとNIST

□ コンテンツを合成・操作する敵対的生成ネットワークまたはその他の技術の機能と出力を検出するため、デジタルメディア企業を含む民間企業との研究機会の実現可能性に関して調査を行うことを義務付け

- 「敵対的生成ネットワークによって生成されたデジタルコンテンツを検出するための革新的なアプローチの実施を通じて、民間企業、NSF、関連連邦機関の間のコミュニケーションと協調を促進・改善しうる、当該機関の政策提言を含む報告書を、共同で議会に提出しなければならない」



2021年国防授權法

- 589F条 軍人とその家族を対象としたサイバー搾取とオンライン詐欺に関する研究
 - (a)研究
 - 本法の制定日から150日以内に、国防長官は以下の調査を完了しなければならない。
 - (1) 軍人とその家族の個人情報とアカウントのサイバー搾取
 - (2) 軍人とその家族を対象とした欺瞞的なオンライン・ターゲティングのリスク。



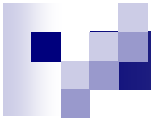
- (b)調査
- (a)項に基づく調査には、次のものを含めるものとする。
 - (1) ～ (2)略
 - (5) 軍人とその家族を対象としたサイバー搾取を行う外国政府および非国家主体が現在もたらしている脅威についての次の情報評価。
 - (A) 当該サイバー搾取が、他の情報戦の手段と比較して、実質的な脅威であるかどうか。
 - (B) 当該サイバー搾取が脅威を増しているかどうか。

- 
- (8) 外国の政府及び非国家主体が、隊員及びその家族を対象に、機械で操作されたメディア（一般に「ディープフェイク」と呼ばれる）を作成又は使用することによってもたらされる脅威についての情報評価。
 - (9) 軍人とその家族のサイバー搾取と欺瞞に対する脆弱性を軽減するための政策決定に関する提言（立法措置または行政措置に関する提言を含む）。

- 
- 軍人とその家族に対する外国政府や非国家主体によるサイバー攻撃やサイバー犯罪は「情報戦 (information warfare)」の一つ
 - 「軍人とその家族を対象としたサイバー搾取を行う外国政府および非国家主体が現在もたらしている脅威」は「情報戦」
 - 軍人とその家族に対するディープフェイクの情報評価
 - 連邦議会の上下両院の軍事委員会に報告書を提出するが機密情報含む



ディープフェイクの検知



大統領令13848

- 2018年9月12日
- 「アメリカ合衆国内の選挙への外国の干渉が発生した場合に一定の制裁を課す大統領令(Executive Order on Imposing Certain Sanctions in the Event of Foreign Interference in a United States Election)」
 - 海外からのサイバー空間を利用した態様を含めた選挙干渉に対し、連邦政府が調査を行うことを義務づけ
 - 干渉が明らかになった場合に経済制裁措置を発動することの規定

- 調査対象

- ☐ 一義的には合衆国選挙(連邦選挙)への外国政府またはその代理による選挙干渉

- 調査の主体

- ☐ 国家情報長官
- ☐ 選挙の後に45日以内に調査を行うことを命じられている

- 選挙インフラストラクチャー

- ☐ 連邦政府若しくは州若しくは地方公共団体により、又はこれらに代わって選挙プロセスを管理するために使用される情報通信技術及びシステム
- ☐ 選挙人登録データベース、投票機、投票集計機器及び選挙結果の安全な伝達のための機器を含む



■ 「外国からの干渉」

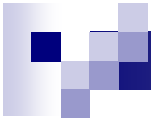
- 「選挙に関し、外国政府又は外国政府の代理人若しくは代理として行動する者の隠ぺいの、詐欺的、欺瞞的若しくは不法な行為又は企てであって、選挙への影響、選挙の結果若しくは報告の結果に対する信頼を傷つけ、若しくは変更し、又は選挙の過程若しくは制度に対する国民の信頼を損なう目的若しくは効果を有するものを含む。」

■ 「外国政府」

- 「合衆国以外の国において、国、州、地方その他の統治当局、政党又は統治当局若しくは政党の職員をいう。」

■ 制裁

- 経済制裁に限定（外交官追放等は含まれない）



SemaFor

■ Semantic Forensics SemaFor

- Michael Spranger, Stefan Schildbach, Florian Heinke, Steffen Grunert and Dirk Labudde, *Semantic Tools for Forensics: A Highly Adaptable Framework*, IMMM 2012 : The Second International Conference on Advances in Information Mining and Management, <https://pdfs.semanticscholar.org/0269/f1bfd314dc7ed416a96cc8d737ac1b8e5c4d.pdf>



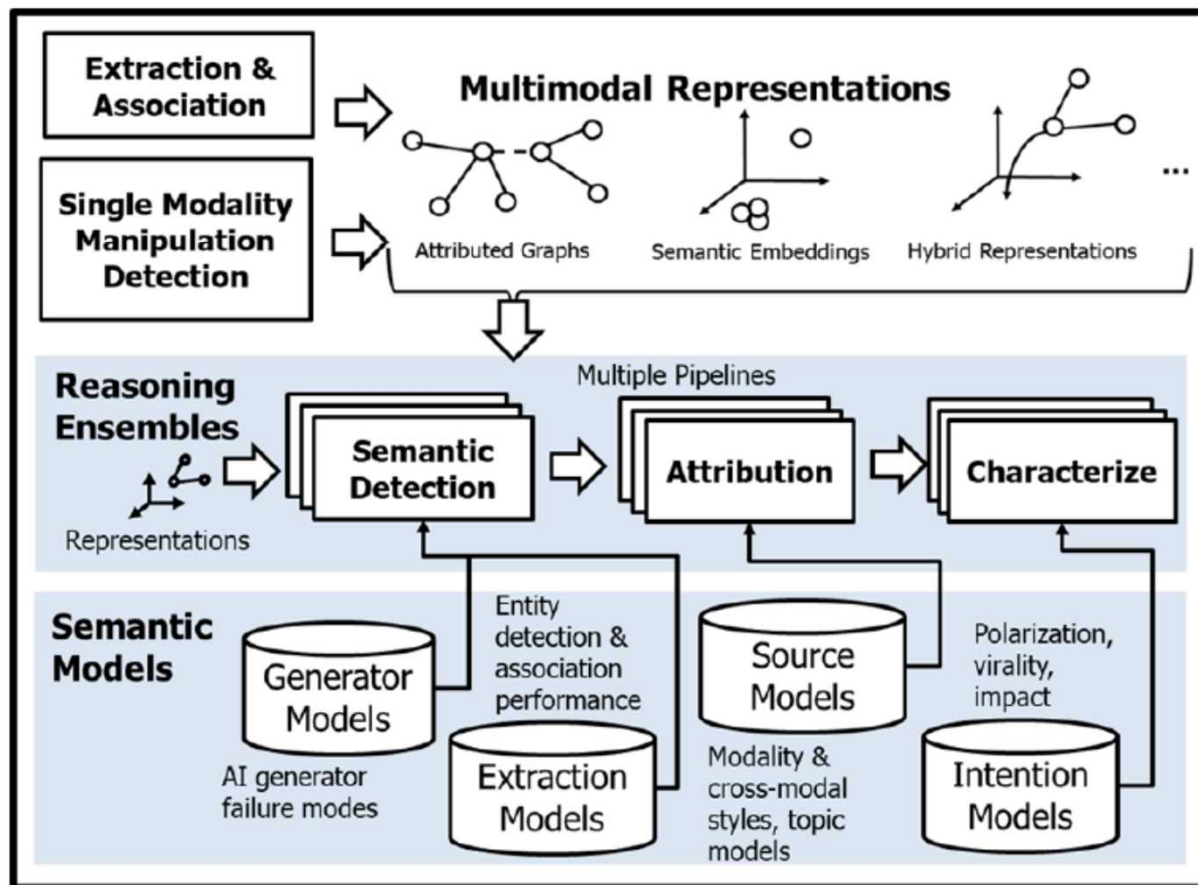
SemaFor

■ DARPA

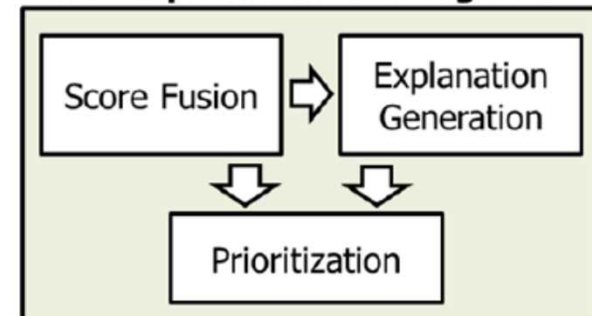
- Semantic Forensics
- SemaFor Project
- Automatic tool for...
 - Detection
 - Attribution
 - Characterization
- Joint study with Locked Martin



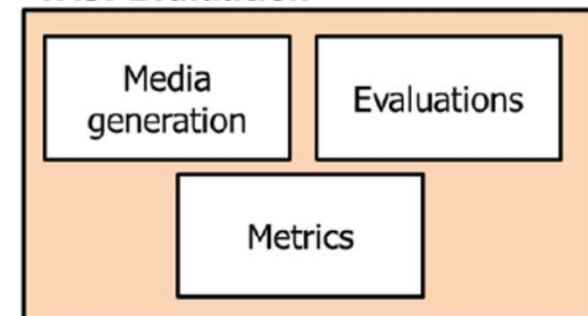
TA1: Detection, Attribution, Characterization



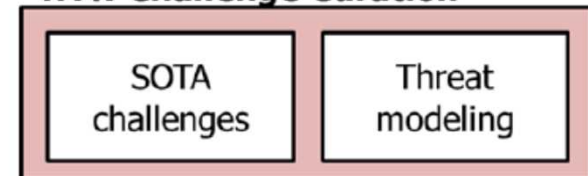
TA2: Explanation & Integration

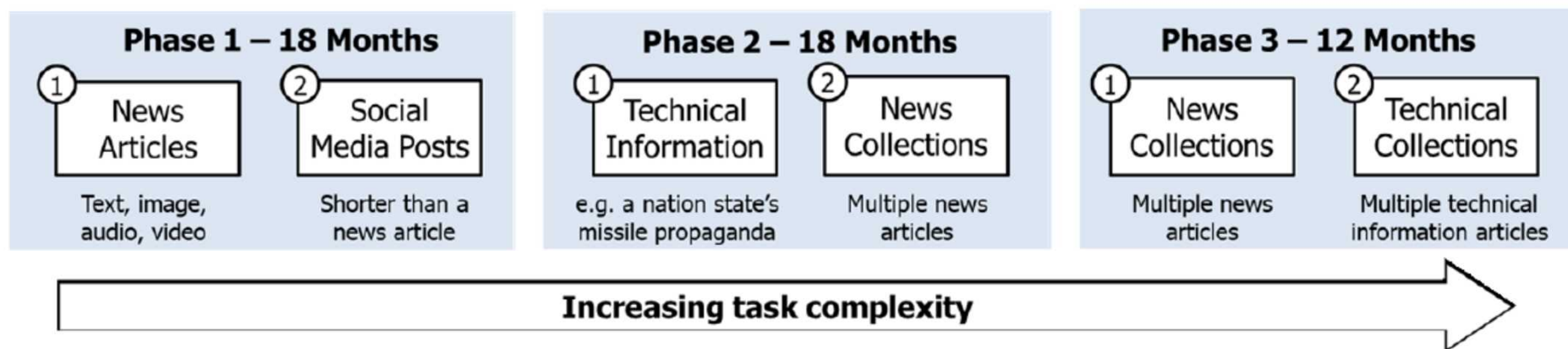


TA3: Evaluation



TA4: Challenge Curation





Quantitative Assessment			Program Goals		
Task	Metrics	Relevant Baselines	P1	P2	P3
Manipulation detection	- Probability of Detection (Pd) - False Alarm Rate (FAR) - Equal Error Rate (EER)	- Human: 60% Pd [Deepfakes] - Image: 80% Pd at 10% FAR / 20% EER - Text entity recognition: 90% F1-score - Audio: 4% EER	80% Pd 10% FAR	85% Pd 8% FAR	90% Pd 5% FAR
Attribution	Pd / FAR	Image: 78% Pd at 10% FAR [camera id]	80% Pd 10% FAR	85% Pd 8% FAR	90% Pd 5% FAR
Prioritization for analyst	Accuracy over degrees of malice	Sentiment analysis: 70-80% F1-score	70% accuracy	80% accuracy	85% accuracy



Semantic Detection

Text (Notional)

NewsWire: April 1, 2019, Bob Smith

On a rainy spring day, a vast, violent group gathered in front of the US Capitol to protest recent cuts in Social Security.

Audio (Notional)



"We'd like to welcome you here on this beautiful spring day. Thank you all for coming out [cheering]..."

Video



Image



Distribution A: Approved for public release. Distribution unlimited.

9

Text (Notional)

NewsWire: April 1, 2019, Bob Smith
On a rainy spring day, a vast, violent group gathered in front of the US Capitol to protest recent cuts in Social Security.

Audio (Notional)



"We'd like to welcome you here on this beautiful spring day. Thank you all for coming out [cheering]..."

Conclusion: Media components consistent across modalities.

Video



"protest"

Image



Distribution A: Approved for public release. Distribution unlimited.

10

<https://www.darpa.mil/attachments/SemanticForensics-IndustryDay-2019-08-12.pdf>



Semantic Detection

Text (Notional)

NewsWire: April 1, 2019, Bob Smith
On a rainy spring day, a vast, **violent group** gathered in front of the US Capitol to protest recent cuts in Social Security.

Audio (Notional)



"We'd like to welcome you here on this beautiful spring day. Thank you all for coming out [cheering]..."

Conclusion: Media components not consistent across modalities.

Image



Video



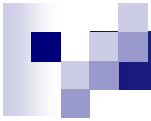
"violent group"

Distribution A: Approved for public release. Distribution unlimited.

11

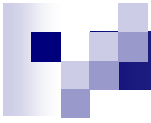


おわりに



ディープフェイクと民意表出

- 公職選挙法の広範な選挙運動規制との矛盾
- 「自由」な民意形成・議席への変換を通じた政治的意思表出と、ディープフェイクを通じた世論誘導
 - 虚偽情報は知る権利に値しない？
 - 「誘導」された民意は自由な民意ではない？
 - 「投げ銭」や広告モデルは営利的表現であり政治的表現ではないから保護に値しない？
 - 海外発のものは保護に値しない？



参考

- 
- Tomoko Nagasako, *Global disinformation campaigns and legal challenges*, INT. CYBERSECUR. LAW REV. (2020) 1:125–136.
 - <https://link.springer.com/article/10.1365%2Fs43439-020-00010-7>
 - 湯淺壱道「郵便・電子投票で民主主義のデジタル化は加速するか」外交66号
 - http://www.gaiko-web.jp/test/wp-content/uploads/2021/03/Vol66_p122-125_postal_and_electric_voting.pdf