

# 偽情報分析に係る技術の開発

(60億円を超えない範囲 / 4年)

## 背景

- サイバー空間の「公共空間化」が進展する中、**事実と異なる加工**や情報の**組み合わせ等**によって、対象の印象やそれに基づく意思決定等を**発信者の意図する方向に導くような情報が存在**。
- このため、文書、画像、動画等の**情報の加工の有無の見極め**、情報発信時点の不自然性や**事実の歪曲等を、エビデンスを踏まえて、一体的、総合的に評価することが必要**。
- 加えて、SNS等のオープンなサイバー空間における偽情報への対処の判断には、社会的影響等、その情報の持つ脅威の評価が重要。
- 今後、**生成AI等の高度化に伴い、サイバー空間における偽情報の増加や複雑化が想定される中**、偽情報の検知からその脅威評価までを一貫して行う、偽情報への対処の判断を支援するシステムが求められている。

## 想定される利用ニーズ

- SNS等のインターネットにおける偽情報は国民生活に影響する可能性があり、マスメディアによる偽情報の打ち消し報道を行う等、対応を判断する必要がある。
- 例えば、特定企業（組織）や製品に関する偽情報は、その企業や製品のイメージダウンにつながり、企業の操業に関して、多大な影響を及ぼす可能性がある。また、災害情報、国内外の事件等についても、意図的に情報が作成された場合、国民生活や経済活動に大きな影響を与えるおそれがある。
- 本構想において開発する技術は、こうした偽情報への対応を迅速に行うための支援のツールとしての活用が考えられる。

## 研究開発の内容

### (1) 要件定義

偽情報分析に係る最新の技術動向や、公的機関や民間企業によるユースケースを調査し、対象とするユースケースを特定するとともに、(2)~(4)の要件定義を行う。加えて、国際会議等において情報収集を行い、海外も含め可能な限り最新の動向を把握する。

### (2) 偽情報検知技術

情報を構成する文章、画像、動画等における加工を総合的に分析する技術を開発する。加えて、情報に付随するメタ情報（時刻情報や地域情報等）に対する客観的な事実の紐づけや発信者の正当性等によるエンドースメントを作成し、情報の真偽性判定に資する要素技術を開発する。

### (3) 偽情報評価技術

情報の受け取り手の反応や発信元の情報等を収集・分析することにより、当該情報に対する社会的影響等を定量化及び可視化し、脅威評価に資する指標・要素を開発する。

### (4) 偽情報検知/評価システム化技術

SNS等のオープンソースを対象とした情報収集方式及び偽情報分析対象の情報の絞り込み手法、(2)及び(3)の技術との連携手法の検討を行い、偽情報検知/評価システムを構築する。

## 想定スケジュール

テーマ	2024年	2025年	2026年	2027年
(1)	ユースケース、技術動向の調査、要件定義の検討			
(2)	偽情報検知手法の確立	中間評価 (ステージゲート)	開発・検証	事後評価
(3)	偽情報評価手法の確立	中間評価 (ステージゲート)	開発・検証	事後評価
(4)	偽情報検知/評価システムの仕様確定	中間評価 (ステージゲート)	システムの構築・検証	事後評価