

2022年3月3日

情報処理学会 第84回全国大会 ～変わる社会と情報処理～

---

# Residential IP Proxyサービスを用いた 位置情報・ターゲット広告の調査

明治大学 総合数理学部

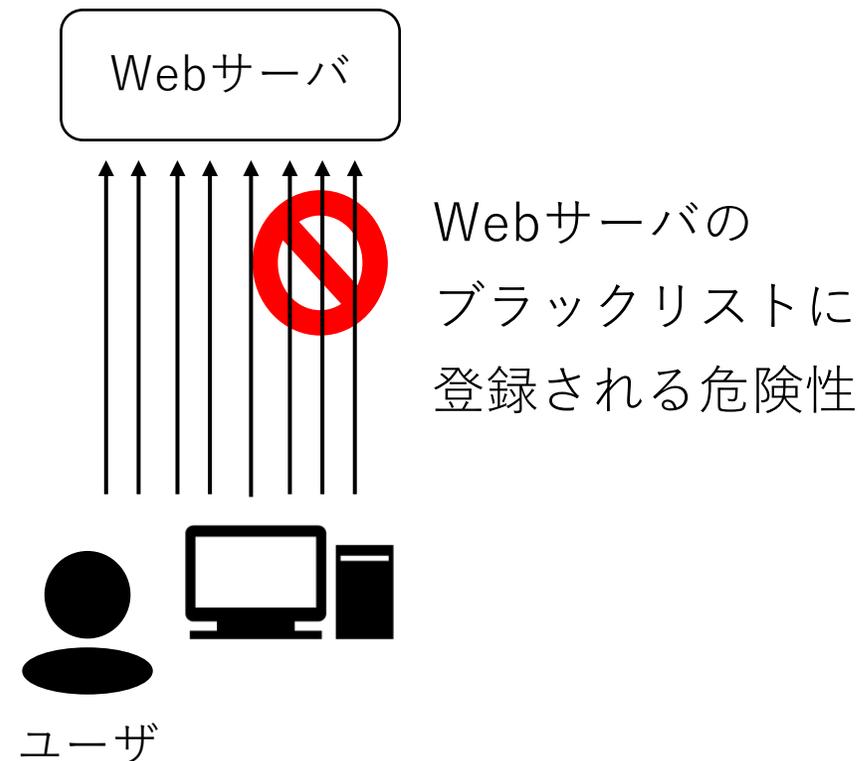
福田 ひかり, 井窪 竜矢<sup>○</sup>, 菊池 浩明

# 背景：Residential IP Proxy (RESIP)の必要性

- ネット検閲



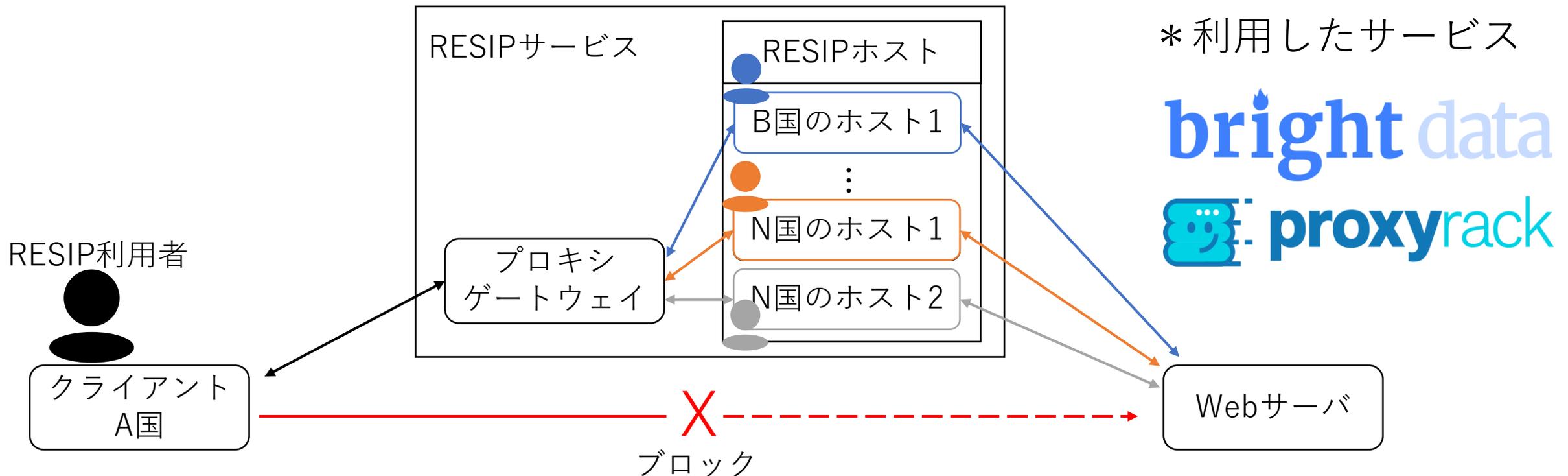
- 同一サイトでの情報収集



# 背景：RESIPとは

プロキシ：クライアントとサーバ間の通信を中継するサービス

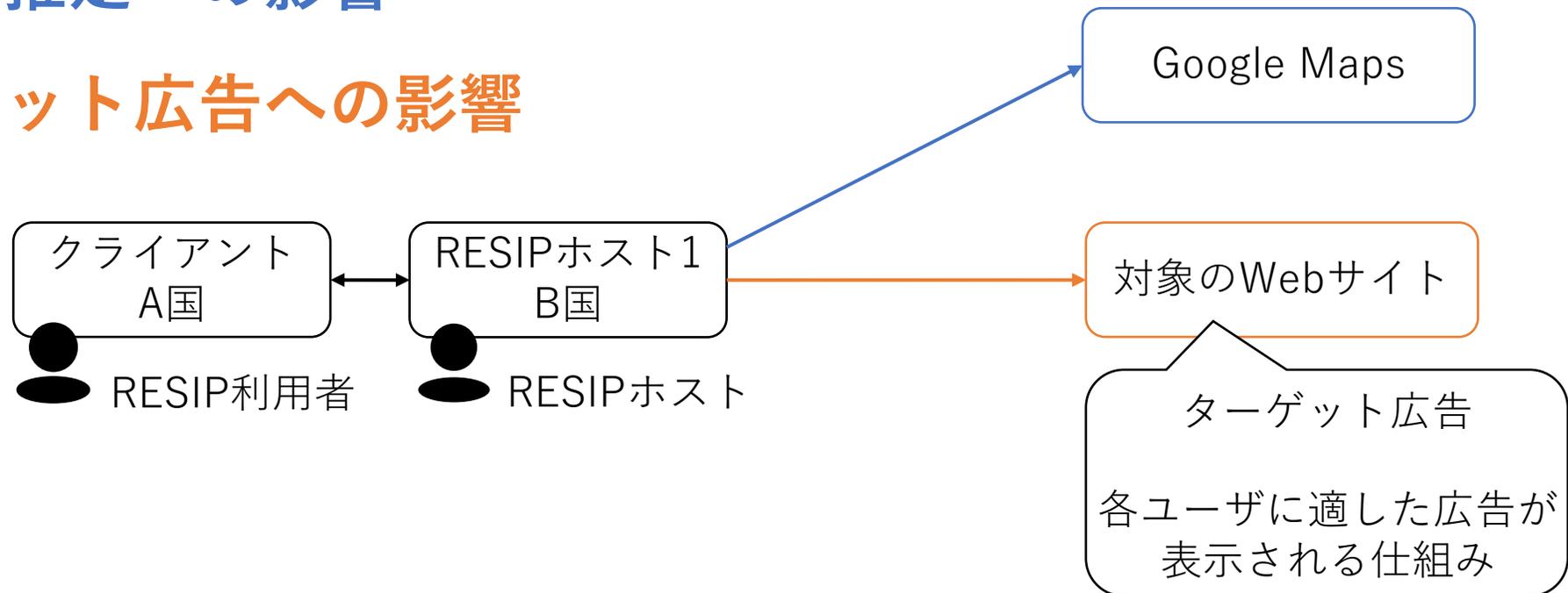
RESIP：中継機を一般の住宅から承諾を得て収集しているプロキシ



# 研究目的：RESIP利用者の調査

目的：RESIPサービスを用いることによる影響を明らかにする

- 現在地推定への影響
- ターゲット広告への影響

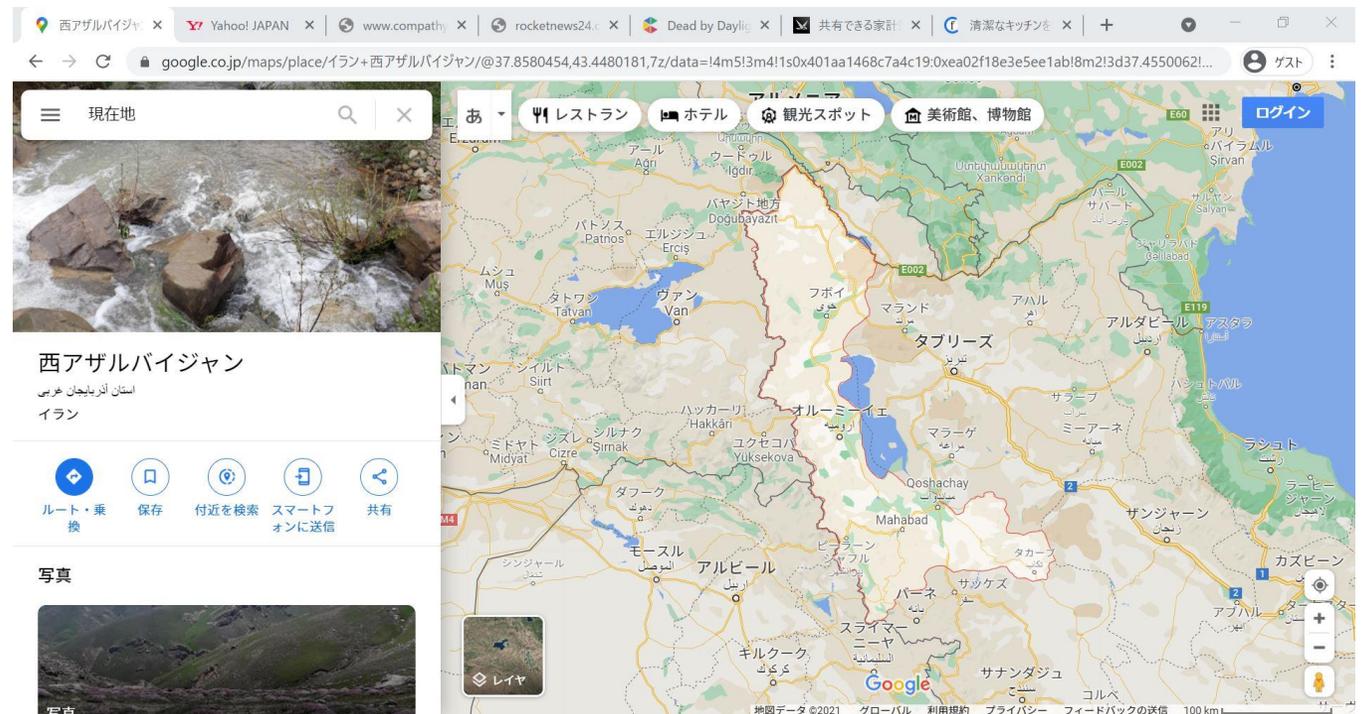


# 実験 1 : 位置情報

## Google Mapsの現在地情報が変化するかの調査

- Bright Data
- ProxyRack

それぞれ40個ずつの  
RESIPホストを対象に  
調査する



現在地がイランと表示された例

# 実験2：ターゲット広告

## 実験手順

1. RESIPを経由する
2. 対象のWebサイトにアクセス
3. サイトに表示された広告数を記録
  - ターゲット広告数
  - RESIPの影響を受けた広告数
4. 5つのWebサイトを対象に手順3を実行
5. 40個のRESIPを対象に手順2, 3, 4を実行



Webサイト：folkの例

対象のWebサイト	Yahoo! JAPAN	神ゲー攻略	folk	ロケットニュース 24	Smartlog
-----------	--------------	-------	------	-------------	----------

# 実験3：実験2の一部を自動化

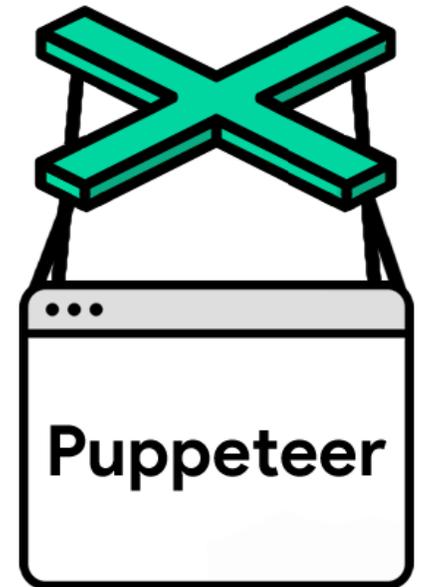
## 実験手順

～自動化～

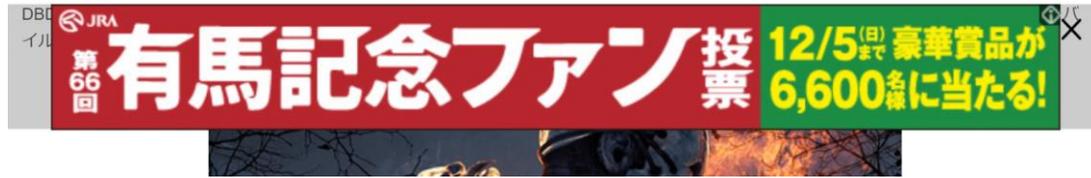
1. RESIPを経由する
2. 対象のWebサイトにアクセス
3. アクセスしたサイトの画像を保存
4. 5つのWebサイトを対象に手順3を実行
5. 100個のRESIPを対象に手順2, 3, 4を実行

～手動～

6. 手順3で保存した画像の広告数を記録



# 影響を受けたとみなす条件



影響を受けていない



条件2：日本語表記で**海外在住者向け**の内容



条件1：**言語**が変化した



条件3：日本語表記で**現在地と異なる地名**が入っている  
大田区にあるIPアドレスを使用した際に表示された広告

- 条件1～3の**いずれか**に当てはまる場合に影響を受けたとみなす

# 実験1：結果

	Bright Data	ProxyRack
現在地の変化	0(個)/40(個)	38(個)/40(個)

- Bright Dataは「現在地を特定できませんでした」というGoogle Mapsのエラーが出力された



現在地を取得できなかった例

国名	個数
韓国	13
日本	7
アメリカ	4
ロシア	3
カナダ	2
インドネシア	2
ウクライナ	1
ベトナム	1
イラン	1
スペイン	1
南アフリカ	1
ナイジェリア	1
香港	1
計	38

- ProxyRackは現在地の変化が生じた
- 現在地として表示された地域は、任意に接続されたIPアドレスの所在地に依存する

# 実験2, 3 : 結果

- 手動における広告が変化した割合は平均**62%**であった
- yahooは自動ではターゲット広告を観測できなかった
- プログラムを用いた自動化の結果は手動に比べ**49%低い**

ターゲット広告が変化した割合

	Bright Data		ProxyRack		平均	
	手動	自動	手動	自動	手動	自動
yahoo	0.05	×	0.00	×	0.03	×
game	0.74	0.02	0.48	0.21	0.65	0.10
folk	0.82	0.02	0.68	<b>0.43</b>	<b>0.74</b>	0.19
rocket	0.54	<b>0.05</b>	<b>0.70</b>	0.12	0.61	0.09
smart	<b>0.83</b>	0.02	0.52	0.40	0.68	<b>0.20</b>
平均	0.68	0.02	0.55	0.27	<b>0.62</b>	0.13

# 結論

---

## 位置情報

- ProxyRackでは40個中38個RESIPホストで変化が確認できた
- 位置情報が取得できないRESIPサービスもある (Ex. Bright Data)

## ターゲット広告

- 手動ではターゲット広告の平均62%がRESIPの影響を受けた

## 今後の課題

RESIPホスト側がRESIP利用者による影響を受けるか調査する