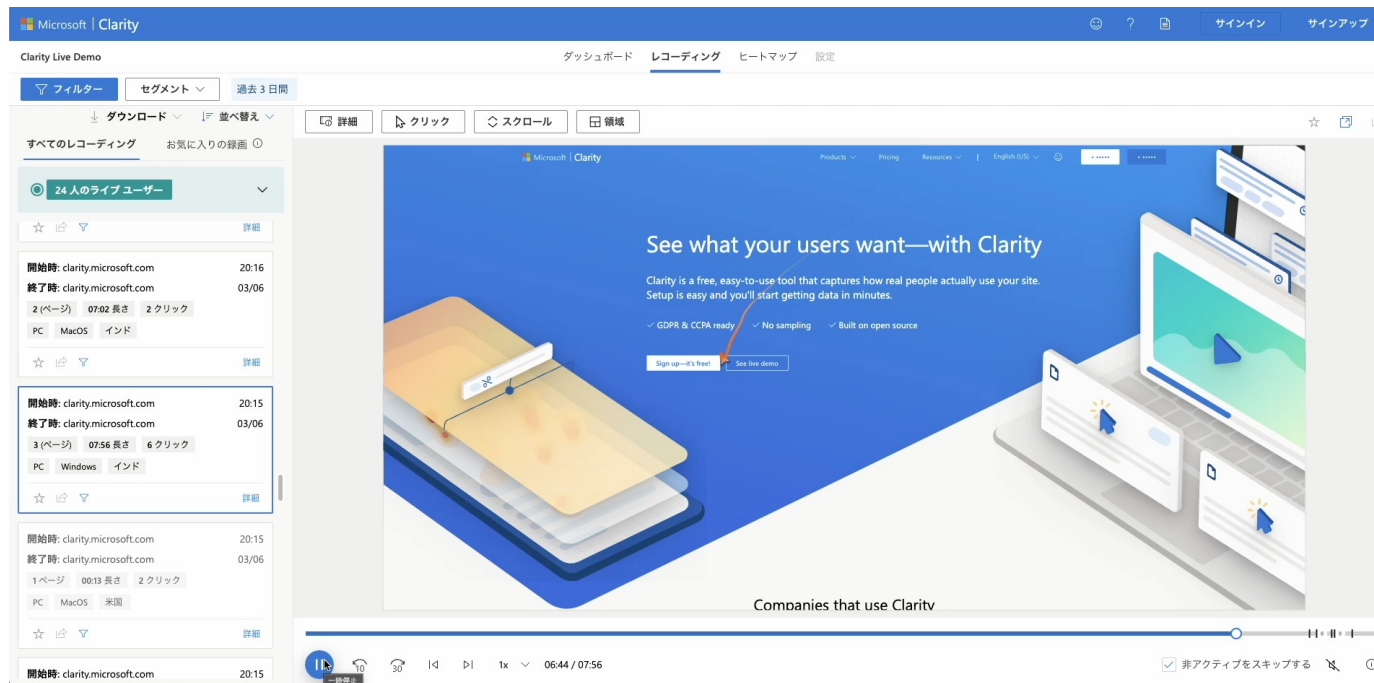


プライバシーポリシーに記載のない  
隠しセッションリプレイサービスの  
導入状況調査

---

菊池研 梶間大地

# セッションリプレイサービスとは



Microsoft Clarityデモ

ウェブサイトの改善ツール

個々のユーザの操作を記録

利用目的

- UX改善
- エラー特定

# 研究背景

入力内容, マウス, キーボード操作をセッションリプレイサービス  
サーバ宛に送信している



構造によってユーザにとって機微な情報が送信されてしまう

そうでなくてもこれらの情報を取得されることに否定的な感情を持つ  
ユーザもいると考えられる

```
▼ 294: {Kind: 18, Args: [762, "k", true, true], When: 4602}
▶ Args: [762, "k", true, true]
  Kind: 18
  When: 4602
▼ 295: {Kind: 18, Args: [762, "か", true, true], When: 4797}
▶ Args: [762, "か", true, true]
  Kind: 18
  When: 4797
▼ 296: {Kind: 18, Args: [762, "かj", true, true], When: 4848}
▶ Args: [762, "かj", true, true]
  Kind: 18
  When: 4848
▼ 297: {Kind: 18, Args: [762, "かじ", true, true], When: 4921}
▶ Args: [762, "かじ", true, true]
  Kind: 18
  When: 4921
▼ 298: {Kind: 18, Args: [762, "かじm", true, true], When: 5133}
▶ Args: [762, "かじm", true, true]
  Kind: 18
  When: 5133
▼ 299: {Kind: 18, Args: [762, "梶間", true, true], When: 5147}
▶ Args: [762, "梶間", true, true]
  Kind: 18
  When: 5147
```

mouseflow導入サイトでの  
テキスト入力時のリクエストの一部

# 先行研究

---

- ウェブサイトのプライバシーに関する研究
  1. サードパーティのスクリプトによる情報流出について50,000サイトを調査[1]
    - 6つのセッションリプレイサービスのDOMの収集によって**パスワードなどの機密情報**が流出
  2. 病院の19,483のウェブサイトのセキュリティとプライバシーを調査[2]
    - **ユーザ名, 電話番号, メールアドレス**をセッションリプレイサービスに送信するサイトを**確認**
- 操作による個人識別研究
  - マウスの移動パターンによる個人識別[3] → 70%以上の精度で識別
  - コンピュータのスリープ復帰時の操作に着目した研究[4] → 4人ではEER2.5%, 12人では10.0%で識別
  - キーボードとマウス操作ログから識別する研究[5] → ログの操作者が同一であることを90%以上で識別
  - スマートフォンで文書閲覧時の操作から識別[6] → 複数の特徴量を用いることで識別が可能

[1]. Gunes Acar et al. “No boundaries: data exfiltration by third parties embedded on web pages”, Proceeding of the 20th Privacy Enhancing Technologies Symposium, pp.220-238, 2020.

[2]. Xiufen Yu et al. “Got Sick and Tracked: Privacy Analysis of Hospital Websites”, IEEE European Symposium on Security and Privacy Workshops, pp.278-286, 2022.

[3]. 泉正夫, 長尾若, 宮本貴朗, 福永邦夫, “マウス操作の特徴を用いた個人識別システム”, 電子情報通信学会論文誌B, pp.305-308, 2004.

[4]. 須田 恭平, 石田 繁巳, 稲村 浩, 中村 嘉隆, “日常的な家電操作による人物識別のためのマウス操作による検討. 情報処理学会研究報告, モバイルコンピューティングと新社会システム研究会”, pp.1-6, 2021.

[5]. 木村悠生, 猪俣敦夫, 上原哲太郎, “深層学習を用いたキーボード入力とマウス操作情報による個人識別”, コンピュータセキュリティシンポジウム (CSS2022), pp.493-499, 2022.

[6]. 渡邊裕司, 市川俊太, “スマートフォンにおけるタッチ操作の特徴を用いた継続的な個人識別システムの検討”, コンピュータセキュリティシンポジウム (CSS2012), pp.797-804, 2012.

# 研究目的

---

1. セッションリプレイサービスが取得する情報に対するユーザ感情の調査
2. セッションリプレイサービスによる個人識別性の調査
3. セッションリプレイサービスの導入状況を明らかにする
4. 導入サイトの傾向を明らかにする
5. 導入サイトがユーザに対して適切に公表を行っているかを確認する



セッションリプレイサービス，事業者は，ユーザのプライバシーを守っているか

# 取得される情報に対するユーザ感情の調査

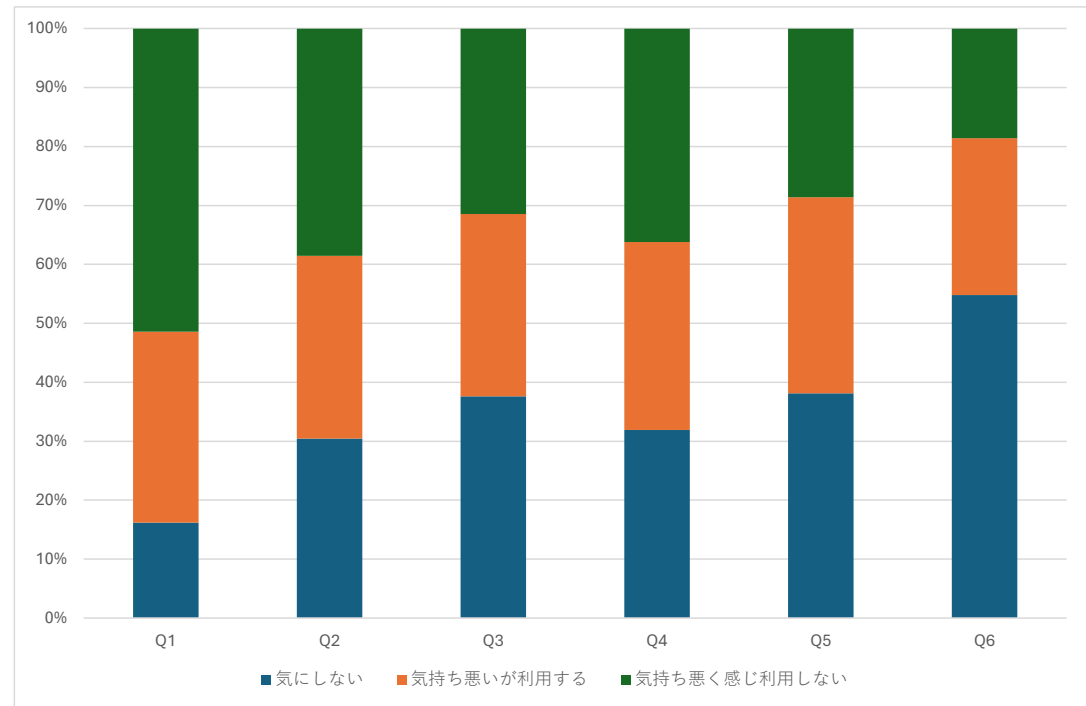
- 方法：クラウドソーシングでアンケート調査
- 参加者：210人
- アンケート内容
  1. 参加者属性について
  2. 架空のサービスに対しての印象
  3. 参加者のセキュリティ意識
  4. セッションリプレイサービスの認知

	属性	人数
年代	20代以下	18
	30代	84
	40代	71
	50代	30
	60代以上	7
性別	男性	98
	女性	112
ウェブサイト 閲覧頻度	毎日	199
	週5~6日	10
	週3~4日	1
	週2以下	0

参加者属性

# 架空のサービスに対する印象の調査結果

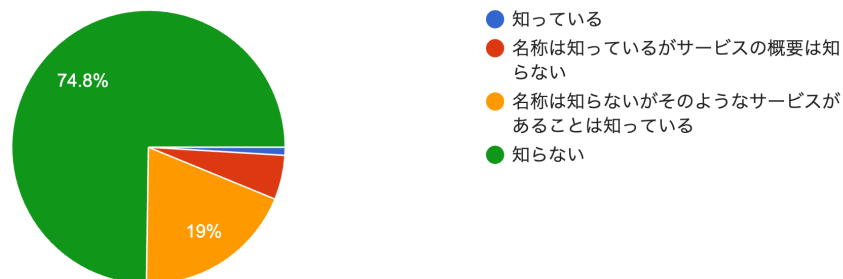
質問文	
1	あなたのスクロール操作が全て記録されるウェブサービス
2	あなたのマウスの動きから悩んでいる商品を推測し、類似したおすすめの商品を推薦するショッピングサイト
3	あなたの閲覧時間が記録されているウェブサイト
4	あなたのキーボード入力内容から最適な広告を配信するウェブサイト
5	あなたがそのページを何回閲覧したかを記録しているウェブサイト
6	あなたのマウスの動きからページのデザインを改善しているウェブサイト



- 取得されることに対して否定的な傾向
- 情報の利用目的に応じて印象が変わる

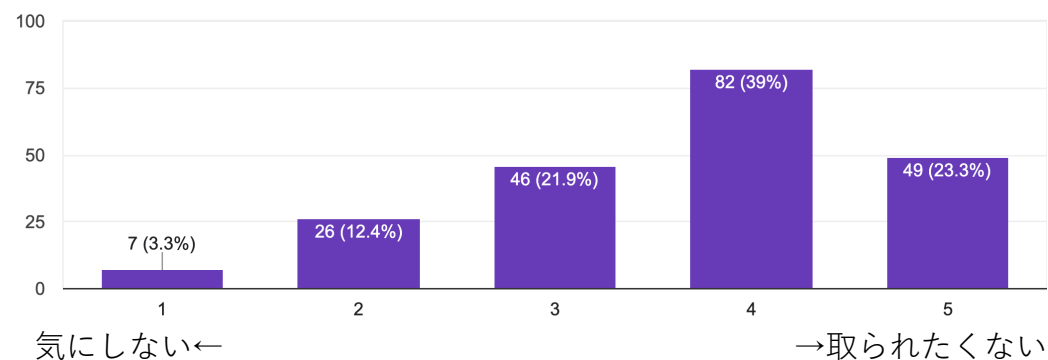
# セッションリプレイサービスの認知調査の結果①

あなたはセッションリプレイサービスを知っていましたか  
210件の回答



- 知っているユーザは2人(1%)
- **157人(74.8%)**がこのようなサービスがあることを知らない

マウス操作やキーボード操作が取得されることに対してどのように感じるかお聞かせください  
210件の回答



- **131人(62.3%)**が取得されることに対して否定的な回答

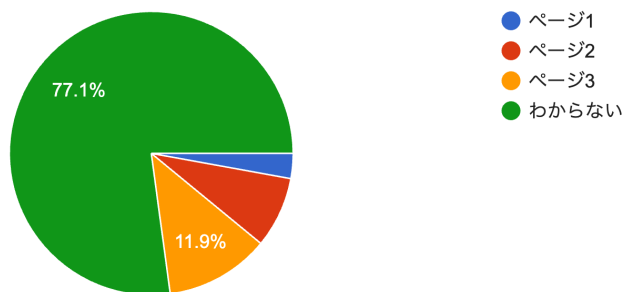


## セッションリプレイサービスの認知調査の結果②

ユーザがセッションリプレイサービスの利用に気づくことができるのかを調査

1. 同一のウェブサイトを3つ用意
2. 内1つにマウスとスクロールの動きをサーバに送信するプログラムを設置（ページ2）
3. 参加者にアクセスしてもらいどのページでセッションリプレイサービスを利用していたかアンケート

下記の3つのウェブページのうち、1つのサイトではセッションリ... ページ1 2. ページ2 3. ページ3  
210件の回答



**17人(8.1%)**がページ2を選択

→ほとんどのユーザは利用されていても気づかない

# 描出操作による個人識別

マウスでの「☆」描出操作による識別

- 研究室の学生15人×5回を収集
- 分析手法：ランダムフォレスト

## 使用した特徴量

マウス操作に要した時間

カーソル速度の平均, 中央値, 標準偏差

軌跡の $x$ 軸の最大値, 最小値

軌跡の $y$ 軸の最大値, 最小値

描出した軌跡の $x$ 軸の最大値, 最小値

描出した軌跡の $y$ 軸の最大値, 最小値

総移動距離

データ取得数



実験に使用したウェブサイト

## 描出操作による個人識別結果

---

	値
Accuracy	0.3697
Precision	0.4333
Recall	0.3502
F <sub>1</sub>	0.3407

- **約30%**で識別可能
- 先行研究と比較して精度が低い
- 取得するデータ数を増やすことで精度が上がる可能性

# ブラウジング操作による個人識別

## ブラウジング操作による識別

- 参加者：研究室の学生15人
- タスク3回を行なった際のマウスとスクロールのログを取得
- 分析手法：ランダムフォレスト

### 使用した特徴量

マウス操作に要した時間

カーソル速度の平均, 中央値, 標準偏差

スクロール速度の平均, 中央値, 標準偏差

軌跡の $x$ 軸の最大値, 最小値

軌跡の $y$ 軸の最大値, 最小値

総移動距離



使用した研究室サイト

# ブラウジング操作による個人識別結果

---

	値
Accuracy	0.3618
Precision	0.2786
Recall	0.3092
F <sub>1</sub>	0.2681

**約30%**で識別可能

描出操作同様，先行研究と比較して精度が低い

取得するデータ数を増やすことやフィンガープリントを利用することで精度が上がる可能性

# サービスの普及率調査

サービス名	検索ワード	検索ヒット数
Microsoft Clarity	clarity.ms	91,110
Hotjar	hotjar.com	293,807
Mouseflow	mouseflow.com	27,641
Yandex	yandex.ru	1,000,000
ContentSquare	contentsquare	1,670
CrazyEgg	crazyegg.com	28,186
Dynatrace	dynatrace.com	6,924
foresee	foresee.com	254
fullstory	fullstory.com	21,132
glassbox	glassbox	806
inspectlet	inspectlet.com	10,340
logrocket	logrocket	5,997
luckyorange	luckyorange	13,642
Smartlook	smartlook.com	33,014

PublicWWW（HTMLソースコード内を含めて検索できるサービス）を利用

調査するサービスの選定のみに利用

- コメントアウトされたソースコード
- 実際のソースコードと検索結果のずれ

# 導入状況の調査

- 対象

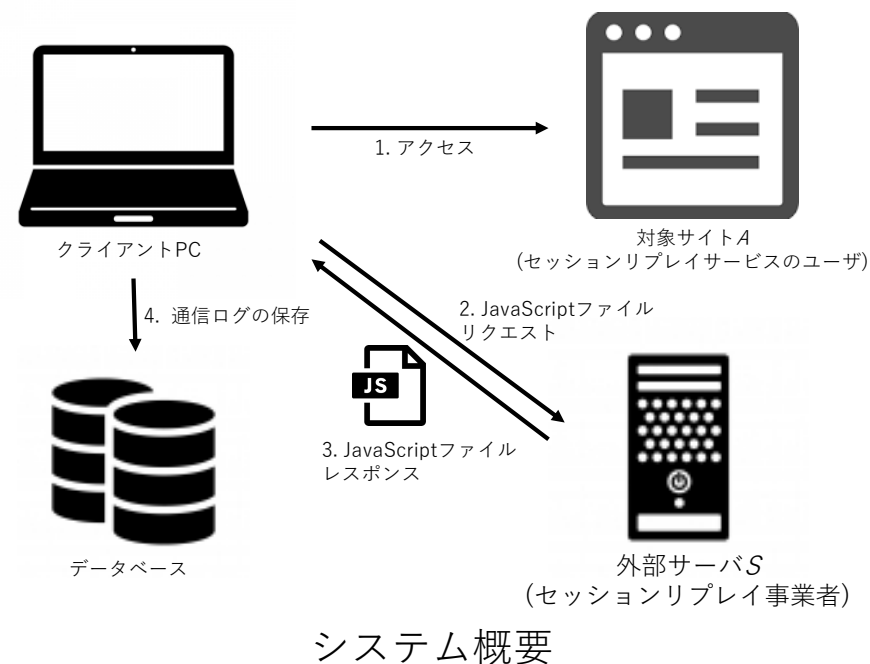
Tranco(トップサイトランキング)から

国内 (.jpドメイン) : 11,523サイト

国外 : 17,582サイト

- 調査方法

1. ブラウザ自動化ツールSeleniumでサイトAにアクセス
2. アクセス時に外部サーバSに送信されるHTTPリクエストをデータベースに格納
3. 既知のセッションリプレイサービスのURLからスクリプトを読み込んでいるサイトを特定



# 国内の導入状況結果

サービスごとの導入数

サービス	導入数
<b>Microsoft Clarity</b>	<b>702</b>
Hotjar	89
Mouseflow	68
Yandex	4
ContentSquare	20
CrazyEgg	40
Dynatrace	2
foresee	1
fullstory	6
glassbox	2
inspectlet	0
logrocket	0
luckyorange	7
Smartlook	2
合計	<b>943</b>

業種ごとの導入数

業種分類	導入数
<b>情報通信</b>	<b>346</b>
<b>卸売, 小売</b>	<b>192</b>
製造	72
宿泊, 飲食サービス	30
教育, 学習支援	29
生活関連サービス, 娯楽	25
学術研究, 専門・技術サービス	40
サービス	91
金融, 保険	20
建設	10
電気・ガス・熱供給・水道	5
医療, 福祉	9
不動産, 物品賃貸	21
運輸, 郵便	8
複合サービス	1
不明	44
合計	943

国内の導入は**943**/11,523サイト (**8.2%**)

Microsoft Clarityが大部分を占める

情報通信業：**346**/943(**37%**)

卸売、小売業：**192**/943(**20%**)

が多い傾向



## 国外の導入状況

サービス	導入数
Microsoft Clarity	818
Hotjar	1127
Mouseflow	86
Yandex	791
ContentSquare	88
Crazyegg	380
Dynatrace	39
foresee	35
fullstory	110
glassbox	10
inspectlet	31
logrocket	6
luckyorange	33
Smartlook	21
合計	3575

3575/17582(20.3%)

→国内(8.2%)より導入割合が多い

国内と比較して様々なサービスが利用されている

# 導入サイトのカテゴリ比較

国内		国外	
カテゴリ	数	カテゴリ	数
ビジネス/経済	181	テクノロジー/インターネット	857
テクノロジー/インターネット	150	ビジネス/経済	481
買い物	130	教育	441
未分類	60	ニュース	258
健康	57	買い物	151
旅行	41	健康	134
娯楽	40	金融	128
教育	39	娯楽	85
ニュース	30	Web ホスティング	77
金融	28	未分類	76
仕事探し/キャリア	22	疑わしい	76
スポーツ/レクリエーション	21	政府/法務	73
車両	14	旅行	72
レストラン/食品	13	参考	66
社会/日常生活	11	ゲーム	59
不動産	10	オフィス/ビジネスアプリケーション	51

NortonLifeLock社のNorton Safe Web  
を利用してカテゴリ分け

国内、国外ともに

- テクノロジー/インターネット
- ビジネス/経済

が多い傾向

Norton Safe Webによる導入サイトの上位14カテゴリ

# プライバシーポリシーの調査

- 対象

サービスごとにランダムに50サイト

- 国内：300サイト
- 国外：400サイト

- 調査項目

- 情報の利用目的の記載の有無
- 具体的な利用サービス名

- 方法

- 手動

### (3) 個人情報の利用

- 当社は個人情報を以下の場合に利用します。
  - 本サービスの通知機能として配信されるメールを受け取ったり、サービス利用のためのログイン認証など、ユーザーが本サービスを利用する場合。
  - ユーザーが興味があると思われる情報を掲示するなど、各ユーザーに最適化した情報を本サービス上に表示する場合。
  - 当社のサービスの改善、開発、マーケティングや、お問い合わせ品質を向上させることを目的として、サービスの利用状況を分析する場合。
  - ユーザーまたは第三者に本サービスについて説明するため、ユーザーの統計情報を用いる場合。
  - ユーザーまたは第三者への連絡を行ったり、本人確認、規約違反行為の調査および対応、問い合わせ内容の調査、商品発送手続きなど、個別の処理や業務に必要である場合。
- 当社は上記以外を目的として、個人情報を閲覧、利用することはありません。
- 個人情報の取得方法や主な項目ごとのより具体的な用途については、はてな個人情報用途表をご覧ください。

情報の利用目的の記載のある  
プライバシーポリシー（一部抜粋[1]）

### 【Microsoft Clarityの利用】

- 当社は、当社ウェブページの利用状況を把握するためにアクセス解析ツールMicrosoft Clarity(以下、Clarity)を利用しています。
- Clarityは、Cookieを利用して利用者の情報を収集します。
- Clarityの利用規約に関する説明は、下記URLよりサイトをご確認ください。
- Microsoft プライバシーステートメント:<https://privacy.microsoft.com/ja-jp/privacystatement>

利用サービスの記載のある  
プライバシーポリシー（一部抜粋[2]）

[1]. はてなブログ, <https://policies.hatena.ne.jp/privacypolicy-ja>

[2]. メロンブックス, <https://www.melonbooks.co.jp/guide/privacy.php>

# プライバシーポリシー調査結果

SRS	国内				国外			
	どちらもあり	目的のみ	サービス名のみ	どちらもなし	どちらもあり	目的のみ	サービス名のみ	どちらもなし
Clarity	12	35	0	3	3	44	0	3
Hotjar	6	41	0	3	6	43	0	1
mouseflow	2	29	0	19	8	39	0	3
crazyegg	10	31	0	9	6	44	0	0
contentsquare	2	22	0	3	4	45	0	1
luckyorange	0	14	0	6	0	27	0	0
Fullstory	1	2	0	1	6	42	0	2
Yandex	1	10	0	4	2	5	0	0
dynatrace	0	8	0	0	1	28	0	3
glassbox	2	1	0	0	0	7	0	1
smartlook	0	7	0	7	0	11	0	2
inspectlet	0	3	0	0	0	7	0	0
logrocket	1	0	0	1	0	6	0	0
合計	37	207	0	56	36	348	0	16

目的, サービス名あり

国内: **12.3%**

国外: **9%**

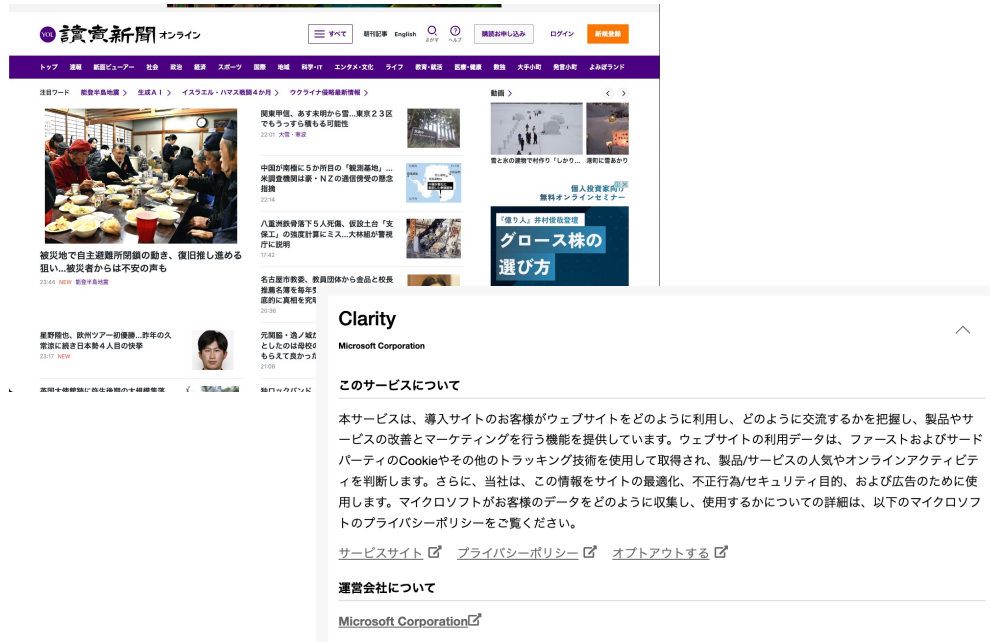
どちらもなし

国内: **18.7%**

国外: **4%**

目的のみを記載しているサイトが多い

# 導入サイトの例



読売新聞(<https://www.yomiuri.co.jp/>)

著名なサイトでも導入が行われている

これらには公表を行っていないサイトも存在する



毎日新聞(<https://mainichi.jp/>)  
公表なし

# Google Analyticsの導入・公表

---

- Google Analytics：ウェブサイトの改善ツール
- セッションリプレイサービスとの違い：個々のユーザの操作を記録しない
- 目的：セッションリプレイサービスの公表が進まない理由を調査

	国内		国外	
	SRS	Google Analytics	SRS	Google Analytics
導入	300	277	400	365
公表	37(12.3%)	<b>116(41.6%)</b>	36(9%)	<b>186(50.9%)</b>

# 考察

## セッションリプレイサービスの公表が進んでいない原因

1. ウェブサイト利用時を想定していない
2. 利用している具体的なサービス名について一部のみを記載

当社では、お客様よりお預かりしました個人情報を以下の目的で利用いたします。

お申込みされた方の個人情報は、コンサルティングサービス実施、ダイレクトメールの送付及びメールマガジン配信のためにのみ利用いたします。また、法令の規定等による場合を除き、お客様の同意を得ずに第三者に提供することはありません。

前述の利用目的達成の範囲内において業務委託する際には、選定基準に基づき個人情報を安全に管理できる委託先を選定した上で当該委託先を適切に監督いたします。

個人情報の提出については、お客様の自由なご判断にお任せいたしますが、必要事項の中でご提出いただけない個人情報がある場合、サービスの1部をお受けいただけない場合はございますのでご了承下さい。

### ■ ダイレクトメールの送付

当社では、お客様より俺お預かりしました住所を、当社よりお送りするダイレクトメールの送付のために使用いたします。ダイレクトメールの送付を希望されない場合は、当社Eメール、電話、またはファックスにて解除お願いいたします。あらかじめご同意の上、お申し込みください。

### ■ メールマガジンの送付

当社では、お客様より俺お預かりしました住所を、当社よりお送りするメールマガジンの送付のために使用いたします。メールマガジンの送付を希望されない場合は、当社Eメール、電話、またはファックスにて解除お願いいたします。あらかじめご同意の上、お申し込みください。

個人情報の利用目的の通知、開示、内容の訂正、追加または削除、利用の停止、商業及び第三者への提供の停止については、下記までご連絡下さい。

## ウェブサイト利用時について記載がない例

### ウェブ解析

弊社は、Google AnalyticsおよびAdobeサービスなどサードパーティのウェブ解析サービスを利用し、サイトの使用状況を追跡・分析し、広告、サイトコンバージョンおよびコミュニケーションの有効性を測定するのに役立っています。これらのサービスプロバイダのツール（クッキー、タグ、ウェブビーコンなど）はこの理解に役立つものです。Adobeによる情報利用を望まない場合は、[ここ](#)をクリックし、Google Analyticsによる情報利用を望まない場合は、[ここ](#)をクリックしてください。

## 一部のサービスのみの公表を行っている例

# 私の提言

---

- 事業者に対して
  1. 取得情報の整理し必要があるものについては正しく公表を行う
  2. セッションリプレイサービスでは個人情報を含む可能性を考慮して通知を行い、サービスを利用していることを公表する
  3. アクセス時に Consent バナーを表示してあらかじめ同意を取得する
- ユーザに対して
  1. サイトで不必要に個人情報を入力しない
  2. プライバシーポリシーで取得されている情報を確認する
  3. 公表なくマウス操作やキーボード操作，入力内容が取得されていることに留意する



# おわりに

---

1. ユーザはマウスやキーボードの情報を取得されることに対して**否定的**
2. セッションリプレイサービスの操作ログによる個人識別は**約30%**
3. セッションリプレイサービス導入状況は国内：**8.2%**， 国外：**20.3%**
4. 導入サイトの傾向
  1. 国内の業種分類は**情報通信業、卸売業、小売業**が多い
  2. 海外においてもテクノロジー/インターネットなどのカテゴリが多い
5. プライバシーポリシーでの公表状況
  1. 目的，サービス名どちらも記載のあるサイトは国内：**12.3%**， 国外：**9%**
  2. 目的もサービス名も記載が行われていないサイトは国内：**18.7%**， 国外：**4%**
  3. 多くのサイトが目的の記載のみを行っている
6. 提言
  - 事業者はユーザの取得情報について細かく公表や同意取得を行う
  - ユーザは公表されていない情報が取得されていることを留意する