日銀・情報セキュリティセミナー

金融分野に求められるユーザブルセキュリティ

2022年7月21日

NTT 社会情報研究所

秋山満昭 akiyama [at] ieee.org

自己紹介：秋山 満昭（Mitsuaki Akiyama）

NTT

- 所属/役職
  - 2007年 日本電信電話株式会社 入社
  - 社会情報研究所 所属
  - 上席特別研究員
- 研究分野
  - サイバーセキュリティ
- 対外的な活動
  - 国内学会の運営委員（MWS, UWS, CSS, CSEC），NISC研究・産学官連携戦略WG, 経産省SC3産学官連携WG, ICT-ISAC, DLPA, SecHack365, …
  - サイバーセキュリティ系の著名学術国際会議プログラム委員（ASIACCS, ESORICS, ACNS等）
- 大学講師等
  - 早稲田大学、大阪大学、岡山大学、横浜国立大学、…
- 書籍
  - 実践Metasploit（2012）
  - コンピュータネットワークセキュリティ（2015）
  - 実践サイバーセキュリティモニタリング（2016）

金融分野に特に関係するサイバー攻撃

フィッシング・スミッシング，
マルウェア感染…

アカウントの不正アクセス，
偽アカウント開設…

不正送金，不正決済…

ユーザブルセキュリティって何？

ユーザビリティ x セキュリティ

ISO 9241-11, JIS Z 8521
「特定のユーザが特定の利用状況において，システム，製品又はサービスを利用する際に，効果，効率及び満足を伴って特定の目標を達成する度合い。」

ISO/IEC27001
「情報の機密性，完全性及び可用性を維持すること。さらに、真正性、責任追跡性、否認防止及び信頼性のような特性を維持することを含めてもよい。」

従来のセキュリティ：システムを中心に考えていた（例：脆弱性のないシステム作り）
ユーザブルセキュリティ：人間（ユーザ）を中心に考える

ユーザブルセキュリティって何？

②よりよい設計・実装・運用を行うためのフィードバック

①ユーザを理解する
「ユーザがどう行動するのか？」「何を考えているのか？」
「どう理解しているのか？」「どのように意思決定するのか？」

ユーザブルセキュリティ研究の始まり

Why Johnny Can’t Encrypt: A Usability Evaluation of PGP 5.0

(USENIX Security 1999)

Johnny：初心者（技術スキルの高くないユーザ）を表す架空の人名

- Eメールの暗号化技術について、なぜユーザがそれを使えないのかに焦点を当てて調査された研究
- セキュリティ（特に暗号化）とユーザビリティの関係について、多くの研究者が強く意識し始めた

Johnnyの苦悩は続く…

Why Johnny can't surf (safely)? Attacks and defenses for web users (Computers & Security)

なぜ安全にWebを利用できないのか？

Why Johnny doesn't use two factor a two-phase usability study of the FIDO U2F security key (FC ’18)

なぜ2FAを利用できないのか？

Why Johnny can‘t opt out: a usability evaluation of tools to limit online behavioral advertising (CHI ’12)

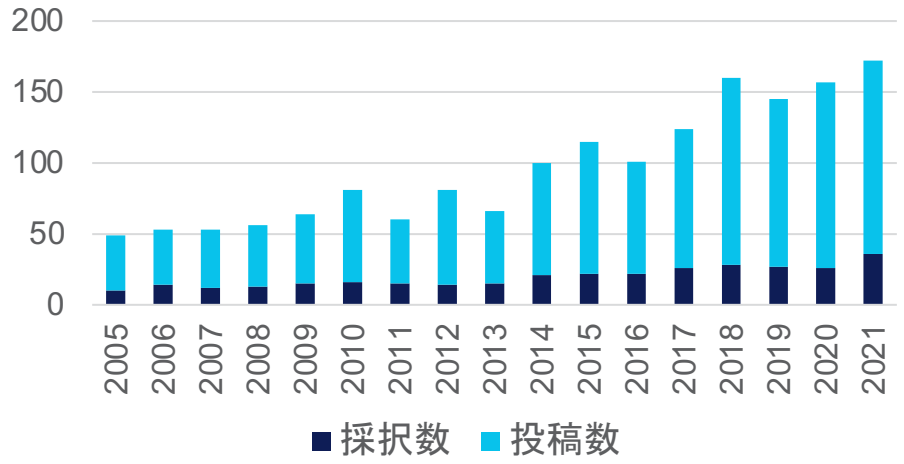
なぜオプトアウトできないのか？

Why Johnny Can't Configure Smart Home? A Behavioural Framework for Smart Home Privacy Configuration (EuroUSEC ‘20)

なぜスマートホームの設定ができないのか？

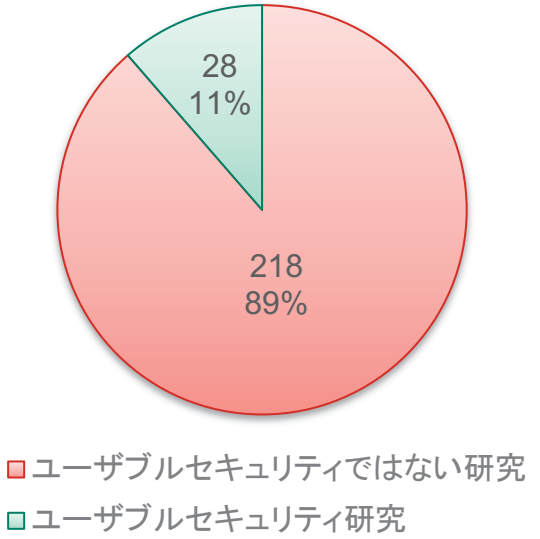
ユーザブルセキュリティ研究の成長

ユーザブルセキュリティの難関国際会議SOUPS



16年で３倍以上の成長

サイバーセキュリティ全般のトップ
国際会議の一つUSENIX Security (2021)



ユーザブルセキュリティ研究が
1割以上を占める

研究対象

ユーザ属性 と **コンテキスト** の組み合わせによって様々な研究対象がある

ユーザ属性		コンテキスト
エンドユーザ	x	ウェブサイト利用
開発者	x	ソフトウェア開発
セキュリティ管理者	x	内部統制
セキュリティエンジニア	x	インシデントレスポンス

研究事例：エンドユーザ × ウェブサイト利用
ブラウザ警告表示の改善（Googleの研究）

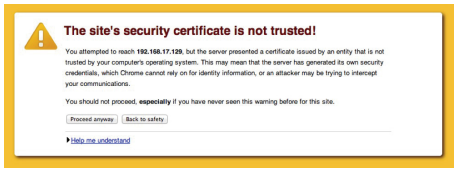
- 何が問題？

Felt et al., Improving SSL Warnings: Comprehension and Adherence, ACM CHI 2015.

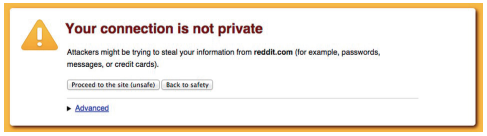
  - ブラウザ上の警告表示（TLS警告）を無視するユーザが多い（内容が理解できない、警告慣れなど）
  - ユーザに状況理解を助け、危険性のあるサイトへのアクセスを防止する効果的な警告表示とは？

- 工夫した警告表示の効果をChromeブラウザのユーザに対して実験
  - ４パターンの警告表示（アイコン、警告文の簡潔さ、色合いを工夫）

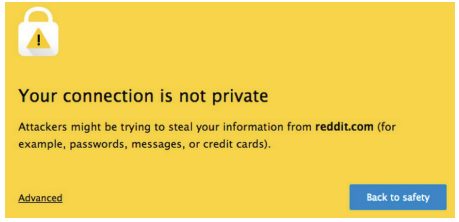
オリジナル (A)



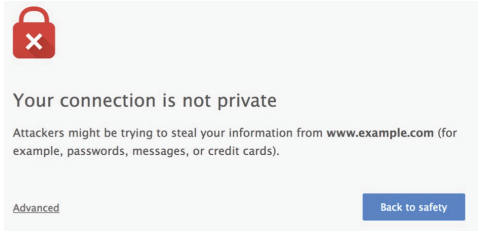
提案 (B)



提案 (C)



提案 (D)

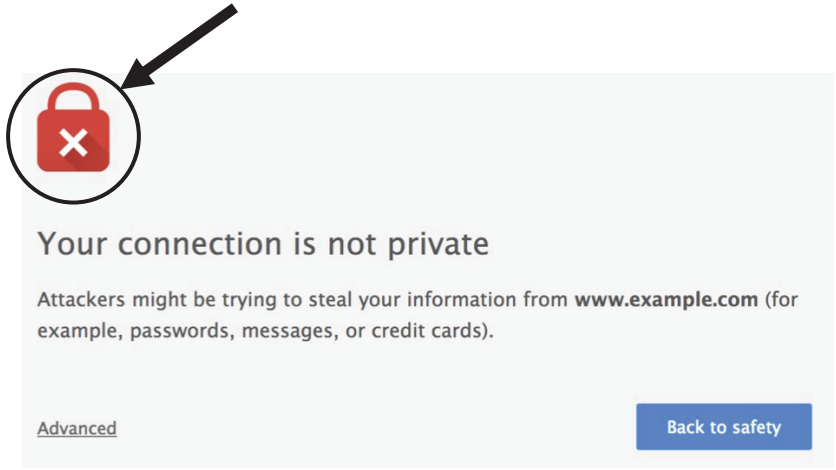


結果：警告に従うユーザは
(A)30.9%, (B)32.1%.
(C) 53.3%, **(D)58.3%**

→ Chromeの実装を(D)に変更

ユーザへの情報提示方法の工夫だけで、セキュリティ技術の性能を向上させることが可能

このインジケータはわかりやすい？



南京錠ではなく、ハンドバッグに見える人もいる…

誰にとっても単純明快でわかりやすいものでなければならない

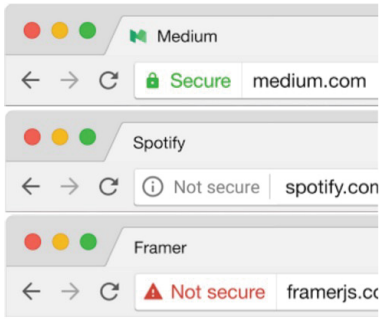
研究事例：エンドユーザ × ウェブサイト利用
セキュリティインジケータの改善（Googleの研究）

この部分

Browser	HTTPS	HTTPS minor error	HTTPS major error	HTTP	EV	Malware
Chrome 48 Win	https://ww	https://mix	https://wro	www.exam	Symantec Co	https://dow
Edge 20 Win	example.	https://mix	wrong.host.bads	example.com	Symantec Co	Unsafe website
Firefox 44 Win	https://www.e	https://mixec	https://expire	www.example	Symantec Corpo	https://spacet
Safari 9 Mac	example.com	mixed.badssl.c	*URL hidden*	example.com	Symantec Cor	downloadgam
Chrome 48 And	https://v	https://mixe	https://v	www.examp	https://v	https://spac
Opera Mini 14 And	www.exam	mixed.badssl.(	wrong.host.ba	www.example	www.syma	*Unavailable*
UC Mini 10 And	Example D	mixed.bad:	*Blocked*	Example D	Endpoint, C	*Blocked*
UC Browser 2 iOS	Example Do.	mixed.bads.	wrong.host.	Example Do.	Endpoint, C.	*Unavailable*
Safari 9 iOS	example.c	mixed.badss	wrong.host	example.con	Symantec	*Unavailable*

乱立するセキュリティインジケータ

- 何が問題？
  - セキュリティの状態を意味するセキュリティインジケータが乱立している
  - 意味の解釈でユーザが混乱している
- ユーザにとってわかりやすいインジケータを検討・実験
  - 色合い・形状・ワードの無数の組み合わせを作り、Chromeユーザで実験（性別・年代・視覚障害などさまざまな属性のユーザを考慮）
  - もっとも意図した動作をユーザが行いやすいものを発見



Felt et al., Rethinking Connection Security Indicators, SOUPS 2016.
https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-porter-felt.pdf

“Usable security is science, …”

（ユーザブルセキュリティは科学である）

Adrienne Porter Felt

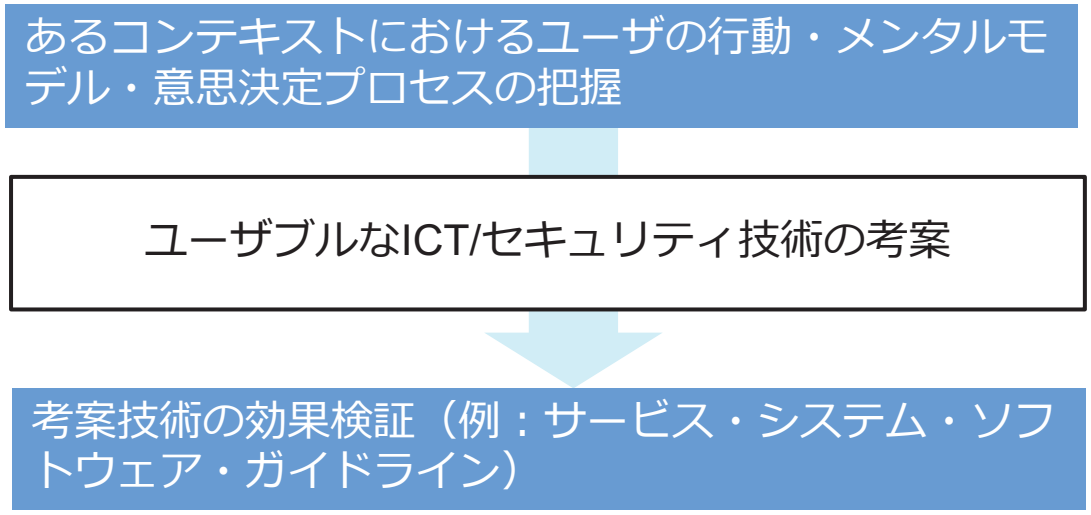
https://www.usenix.org/conference/enigma2016/conference-program/presentation/porter-felt

人間を観測して理解することは難しい、だから科学的に取り組む必要がある

- 正しく観測すること
  - ✕ **実験用デバイス**でブラウザ操作を観測する → 日頃の操作とは異なる
    ○ **普段使いのデバイス**でブラウザ操作を観測する
- 定量的に表せること
  - ✕ **多く**のユーザが□をする → 具体的にどれだけなのかわからない
    ○ **85%**のユーザが□をする
- 再現できること
  - ✕ 同じ条件でも**毎回結果が異なる** → 普遍的な事実ではない
    ○ 同じ条件であれば**毎回結果が同じ**
- 統計的に有意であること
  - ✕ 20代はAを選ぶ人が**多い** → 偶然かもしれない
    ○ 20代がAを選ぶ人が多いのは**統計的に有意**

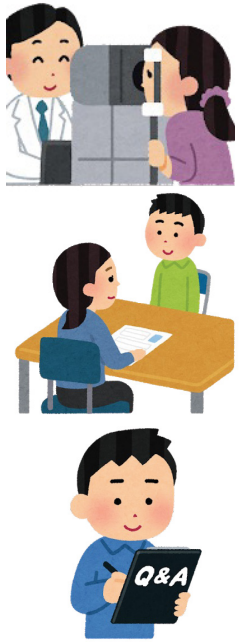
ユーザスタディ

- ユーザに参加してもらって行う調査
  - 実環境/ラボ実験による直接的な観測だけでなく、インタビュー/オンラインサーベイなどの社会科学・心理学的手法なども用いる
  - 問題点の抽出、対策効果を検証
- 目的に応じた
  ２種類のユーザスタディ



ユーザスタディ

- ラボ実験（実験法）
  - 実験室に招き、参加者の行動を観測
- インタビュー（面接法）
  - 対面、電話、オンライン、個人/グループなど、実施形態は様々
- アンケート（質問紙法）
  - アンケートに参加者が回答する
  - **クラウドソーシングサービス**（Amazon Mechanical Turk、Lancersなど）により、**大規模に実施可能**
- その他の方法
  - フィールドワーク/参与観察、ログ収集ツール、オープンデータ活用



オンラインアンケートの例

質問:
この質問はあなたが質問文を読んでいるか確認させていただくためのものです。
どの選択肢も選ばず、次の質問に移ってください。

以下の選択肢の中であなたが一番面白かったと思う
東京オリンピック2020の競技はどれですか?

○  柔道
○  体操
○  卓球
○  スケートボード
○  サーフィン
○  水泳

注意力を試す質問（Instructional Manipulation Check, IMC）でした

オンラインアンケート実施時の注意点

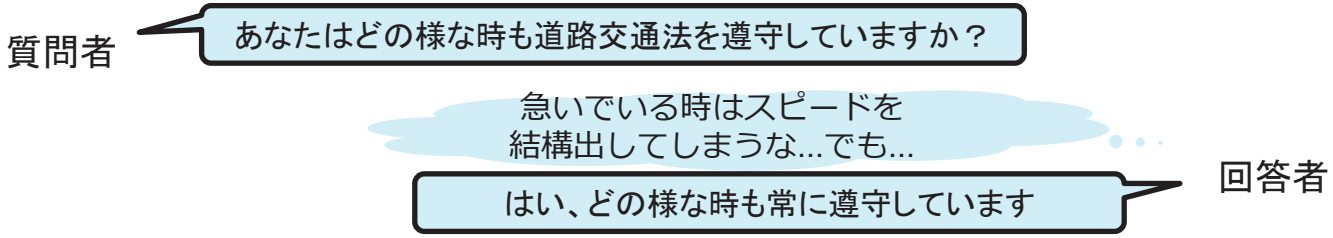
- いい加減に回答する参加者が一定数存在するため、正確なデータを取ることが難しい
  - 参加者は短時間で多くの報酬が欲しいので、急いで回答する
  - アンケートに自動で回答するツールを作成している参加者がいる可能性も
- 対策手法：参加者のスクリーニング
  - 注意力を試す設問（トラップ質問）を設ける
    - 注意せずに回答する参加者を除去
  - 回答時間を測定
    - 回答時間が極端に短い参加者を除去
  - 自由記述を設ける
    - 意図に沿った記述ができていない参加者を除去
      - 例：理由を聞いているのに「OK」「Good」などと回答する場合

ユーザスタディ全般の注意点（1/2）

- 調査環境を一定にする
  - ラボ実験の指示や調査用デバイスを共通化し、参加者毎に差が出ない様にする
    - 悪い例） 参加者AにはPCを利用させ、参加者Bにはスマホを利用させる
    - 悪い例）参加者Aには「自分のペースでやってください」、参加者Bには「なるべく早くしてください」
- 参加者の選定が適切で属性に偏りがない
  - 仮説の母集団に対して代表性のある参加者
    - 悪い例）プロ開発者向けツールを作ったが、参加者が全員大学生
- 生態学的妥当性（Ecological validity）
  - ラボ実験など、研究者が用意した人工的な環境・課題を使って実施する調査の場合、ユーザが日常の活動の場で行なっている実際の行動となるべく関連するようにすること
    - 悪い例）フィッシングメールの騙されやすさを実験する際に、参加者が普段使っていないデバイス（モバイルorPC）・OS・メーラ・ブラウザを利用させる

ユーザスタディ全般の注意点（2/2）

- ホーソン効果（Hawthorne effect）
  - 実験実施者（研究者）の期待に応えようと行動してしまう現象
  - 実験の意図を説明しすぎない様にする
    - 悪い例「実験説明：**操作性を向上させるために改良したツール**について、従来ツールと比較してどちらが操作性が良いか実験します」
    - 良い例「二つのツールを比較してどちらが操作性が良いか実験します」
- 社会的望ましさバイアス（Social desirability bias）
  - “社会的に望ましい行動”を参加者が推察し、実際とは異なる回答/行動をすること



  - 匿名アンケートにするなど、本音の回答を引き出しやすくする工夫をする

ユーザの多様な属性を考える：「普通のユーザ」って誰？

ユーザの属性

- 年齢：若年者～お年寄り
- ジェンダー：男性、女性、LGBT…
- ハンディキャップ：視覚障碍者など
- 母国語：日本語、外国語（日本語以外）

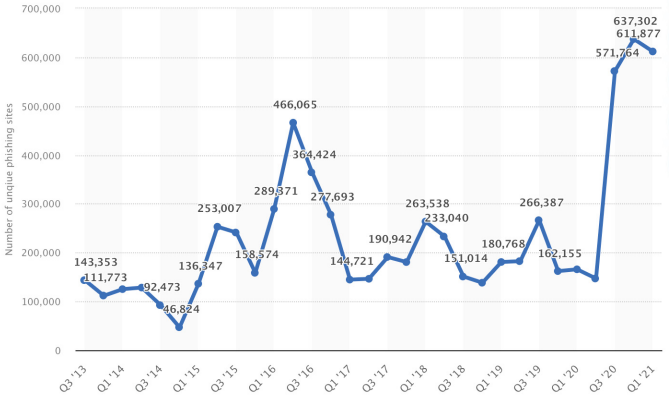
ユーザの属性に応じた設計の検討

- 高い技術レベルを前提とせず、専門用語を避け、平易な説明を行い、必要に応じて詳しく説明する
- 色合いだけでの説明を避ける
- 言語を選べる

ユニバーサルデザイン（汎用的なデザイン）から
インクルーシブデザイン（多様性を許容できるデザイン）へ移行する必要性

フィッシングの増加

- フィッシングは、人間の認知の脆弱性を狙ったサイバー攻撃とも捉えられる
  - システムの脆弱性はパッチ等で修正できるが、人間の認知の脆弱性は対策が容易ではない
  - 人間の騙される原理の理解、サポート技術に関するユーザブルセキュリティの研究が盛んに行われている
- フィッシング対策協議会
  - 手口の詳細が公開されている
    https://www.antiphishing.jp/news/alert/



フィッシングサイト検知数（四半期毎）

https://www.statista.com/statistics/266155/number-of-phishing-domain-names-worldwide/

スミッシング（SMS phishing, Smishing）

- SMSを誘導手段として用いるフィッシングが増加中
- メールでは送信者認証技術（SPF、DKIM、DMARC、BIMI等）によって「なりすましメール」の対策が可能だが、SMSにはそのような送信者認証技術がない
- SMSは情報が少なく、フィッシングを見分けるための手がかりが少ない
- 各種サービスのSMS通知/認証などにおいて、日頃からSMSのリンクをクリックするという習慣付け（habituation）がされている

https://www.antiphishing.jp/news/alert/nttdocomo_20220210.html

ユーザに対して行動変容させるための対策を事業者として検討すべき。
（事業者の通知ではURLの代わりにブックマークやアプリで当該ページへアクセスさせるよう誘導、など）

暗号通貨に関するフィッシングや詐欺

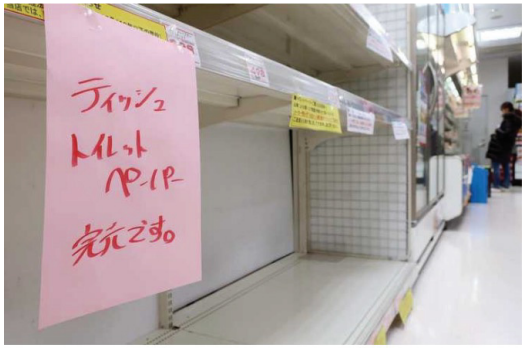
- 偽の取引所サイト/アプリ、ウォレットアプリ、マイニングアプリ、NFT、送金先アドレス…
  - これまでのフィッシングと同様に騙しの手口をベースにしているが、人間の認知を狙っている以上は新しい技術への過渡期においてはかならず発生する
- サンドイッチ攻撃
  - 価格操作を行う攻撃
    Zhou et al., High-Frequency Trading on Decentralized On-Chain Exchanges, IEEE S&P 2021.
  - ユーザは、暗号通貨の仕組みが難しすぎると感じており、いちいち取引価格の確認をしない
    Wang et al., Impact and User Perception of Sandwich Attacks in the DeFi Ecosystem (CHI2022)

フィッシングに対してこれまで行っていたような確認手順（送信元メールアドレス、内容、…）は、暗号通貨の難解な仕組みとインターフェースにおいてはより難しくなっている

誤情報拡散

https://courrier.jp/columns/68892/

https://www.sankei.com/article/20160415-5H2RBURH6NMGFPTWZEIHHIGP2Y/



https://www.sankei.com/article/20200510-GQ4TXA2Z5ZNZHJR5UAPXKT4IFA/

用語の整理

- Misinformation
  - 誤った情報が悪意なく拡散される
- Disinformation
  - 誤った情報が悪意を持って拡散される

誤情報拡散
（いわゆる”フェイクニュース”もこれに含まれる）

誤情報拡散は人々の認知を歪め誤った判断を誘発させる
新しいサイバー攻撃の一種であると捉えることができる

誤情報拡散によって起きた取り付け騒ぎ

- 豊川信用金庫（1973年）：高校生の冗談の会話が発端で拡散
- 佐賀銀行（2003年）：ある人が友人から聞いた誤った情報を別の友人らにメールで送信したのが発端で拡散
- 中国四川省の自貢銀行（2018年）：地元の人物がデマの発端

誤情報拡散の一般的な対策

- ファクトチェック
  - 情報が事実に基づいているかを調査、情報源の確認や複数のデータを突き合わせるなどして、「正しい」「誤り」「根拠不明」などと評価する
  - ファクトチェックサイト：FactCheck.org など
- 警告表示
  - 情報について注意が必要なもの（ または「誤り」「根拠不明」など）について警告を表示
- コンテンツ・モデレーション
  - コンテンツの削除やアカウントの凍結等
    （ただし判断の透明性が重要課題）

一般的には、「ファクトチェックされる」ということや、「情報の信頼性が低い」ことがわかると、ユーザのより慎重な情報発信が期待できる

誤情報拡散の対策の難しさ

- ポスト・トゥルース（post-truth）の時代
  - 真偽は重要ではなく、人々の注目を集めること（面白い、センセーショナル、…）の方が価値が高いと考えること
- バックファイヤー効果（Backfire effect）
  - SNS上で間違いを指摘されると逆に誤情報への信念を強めたり、他人への攻撃性が高まる場合がある
    Mosleh et al., Perverse Downstream Consequences of Debunking: Being Corrected by Another User for Posting False Political News Increases Subsequent Sharing of Low Quality, Partisan, and Toxic Content in a Twitter Field Experiment (CHI2021)
- 誤情報拡散の責任の所在は？
  - プラットフォーム事業者は何をどこまでやればいいか？
  - 総務省「プラットフォームサービスに関する研究会」で議論中
    https://www.soumu.go.jp/main_sosiki/kenkyu/platform_service/index.html

プラットフォーム事業者以外の事業者ができることは？
事業者とユーザとの間での正しい情報を伝えるためのコミュニケーション手段はあるか？

事業者として検討しなければならないこと

- ユーザを理解する方法の確立
- ユーザを混乱させない/負担を軽減する/多様な属性を考える設計
- ユーザとの正しいコミュニケーション手段の確立