最新事例と共に考える サイバーセキュリティ研究倫理

横浜国立大学
大学院環境情報研究院/先端科学高等研究院

吉岡克成

注意：本資料は米国DHSによるメンロレポート[1]とその補足資料[2]を基に、サイバーセキュリティ研究における研究倫理に関する資料作成者(吉岡)の**個人的解釈・意見**をまとめたものです.

本日のおはなし

- メンロレポートとは
- 歴史的経緯
- ベルモントレポートと研究倫理3原則
- 生物医学とICT研究の違い
- メンロレポートが定める研究倫理4原則
- 研究倫理4原則の説明
- 具体的事例で考える

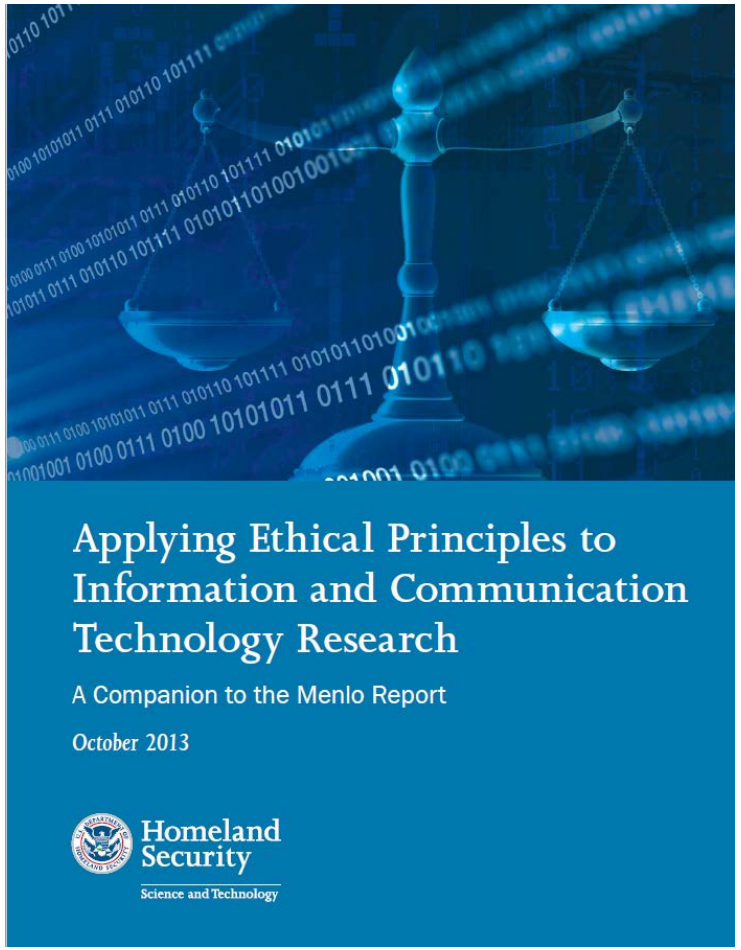
メンロレポート[1]とは

- 2012年8月に米国DHSが発行
- 正式名はThe Menlo Report –Ethical Principles Guiding Information and Communication Technology Researchであり、ICT研究における研究倫理の原則(Principles)を定めるものである†
- 生物医学と行動科学における3原則を定めたベルモントレポートの理念をベースに、さらに1原則を追加し、4原則を定めている
- 15名程度の産学官有識者によるWG構成と執筆

†一部報告書内の説明が不統一であり、Executive SummaryにはThis report proposes a framework for ethical guidelines for computer and information security researchと記載があるものの、他はICT研究に対する言及となっている．実態としてはセキュリティ関連研究が強く意識されている．

補足文書[2]

- Applying Ethical Principles to Information and Communication Technology Research –A Companion to the Menlo Reportというタイトルで2013年10月に米国DHSが発行
- 原則の記述しかないメンロレポートの、より具体的な解釈が記載
- メンロレポートが13ページに対して補足文書は32ページあり、実際の研究事例の説明もある
- メンロレポートの主旨を理解する上ではこちらの補足文書の方が役に立つ(メンロレポート本体のおさらいにもなる)



歴史的経緯

- 1979年に生物医学などの研究における研究倫理を定めた**ベルモントレポート**が制定され、米国政府の研究資金を受ける際にこれに従うことが広く求められるようになる
- 1970年代のDARPANETに始まるインターネットの急速な発展はサイバー攻撃の増加や個人特定できる情報の増加をもたらす.
- 初期のICT(セキュリティ)研究では適切な指標もなく、倫理的検証が不十分なまま行われた. 例えば、マルウェアの取り扱い、サイバー攻撃への反撃、脆弱性への攻撃や公開、機微な情報の収集などである
- これらの研究を経て、ICT研究における倫理問題の重要性とベルモントレポートの考え方を**ICT研究の文脈で解釈**する必要性が生じ、メンロレポートが制定されるに至った

ベルモントレポートが定める
(生物医学における)研究倫理3原則

- **人格の尊重(Respect for Persons)**
  - 研究対象の参加は本人の自由意志によって決まり、インフォームドコンセントによるべきである．本人が意思決定する権利を尊重すること．直接的な研究対象だけでなく、研究によって影響を受ける可能性があるが、自身の意思決定によりこれを決められない個人も保護の対象である
- **恩恵(Beneficence)**
  - 危害を加えないこと．研究により得られ得る恩恵を最大にし、与えうる危害を最小にすること．リスクと危害と恩恵のアセスメントを行うこと．
- **正義(Justice)**
  - 個人は自身の扱いについて平等に配慮を受けるべきであり、研究の恩恵は平等に分配されるべきである．研究対象の選択は公正に行われ、負担は研究対象に対して同等に分担されるべきである．

生物医科学とICT研究の違い

- 規模
  - ベルモントレポートが想定する生物医学では研究者と対象は対面でやり取りをすることが前提であり、対象は数十〜数千人であるのに対して、ICT研究では数百万人規模のデータ収集・分析を行う場合もあり、このような場合に各個人からインフォームド・コンセントを得るのは現実的でない
- 速度
  - 生物医学では多くの研究はマニュアルプロセス(研究室で対面で行われるなど)であり、問題があった場合にもその被害が拡大する前に研究を中止することができる。一方ICT研究では数秒のうちに数百万のデバイスに悪影響を与える可能性があり、リスクと被害について迅速かつ的確な判断が必要である

生物医科学とICT研究の違い

- 情報の集約と相互関係性
  - ICT研究においては情報資源はネットワークを通じて相互接続しており、その関係性が深い。例えば、スマートフォンには友人等の連絡先リスト、名前、住所、電話番号、メールアドレス、ソーシャルメディアアカウント、家族、財務、保険、投資のための口座情報、家庭用機器や自動車を操作するアプリケーションへのアクセス権、個人の写真といった情報が蓄積されている。これにより、デバイスの持ち主だけでなく、それとつながる他人の個人情報の暴露にも繋がる可能性がある
- 非集中化
  - ICTは様々な技術に相互依存しており、テキスト、音声、映像といった通信内容は様々な場所に位置し、様々なエンティティに制御されているため、インフォームドコンセントを得る対象を特定するのが困難と成り得る。

生物医科学とICT研究の違い

- 不透明性
  - 生物医科学では対象と対面するのに対して、ICT研究ではICTを介してその先に人間が多数存在する。直接対面していないため研究が誰にどのように影響を与えるのかを予想するのは難しい。

メンロレポートが定めるICT研究倫理4原則

- **人格の尊重(Respect for Persons)**
  - 研究対象の参加は本人の自由意志によって決まり、インフォームドコンセントによるべきである．本人が意思決定する権利を尊重すること．直接的な研究対象だけでなく、研究によって影響を受ける可能性があるが、自身の意思決定によりこれを決められない個人も保護の対象である
- **恩恵(Beneficence)**
  - 危害を加えないこと．研究により得られ得る恩恵を最大にし、与えうる危害を最小にすること．リスクと危害と恩恵のアセスメントを行うこと．
- **正義(Justice)**
  - 個人は自身の扱いについて平等に配慮を受けるべきであり、研究の恩恵は平等に分配されるべきである．研究対象の選択は公正に行われ、負担は研究対象に対して同等に分担されるべきである．
- **法と公益の尊重(Respect for Law and Public Interest)**
  - 法に従うこと．研究方法と結果の透明性を保つこと．行為に責任をもつこと．

上記のうち最初の３つはベルモントレポートを踏襲し、ICT研究のコンテキストで解釈することで対応している。最後の原則のみは新たに追加されている．

利害関係者(stakeholder)の明確化

- 4原則を論じる上で利害関係者の明確化がまず必要

- **Primary stakeholder:**最終的に影響を受けるエンティティ(例：エンドユーザや製品の購入者)
- **Secondary stakeholder:** 利害を媒介するエンティティ(例：サービスプロバイダ、オペレータなど)
- **Key stakeholder:** プロジェクトの成否に多大な影響を与えうるエンティティ(例：研究者、ベンダ、システム設計実装者、犯罪者、攻撃者)

ICT研究における「人格の尊重」の扱い

- 生物医学での人格の尊重の基本は、影響を受けうる各個人からの「インフォームドコンセント」を得ることである
- ICT研究では、①利害関係者を特定することが難しい、②利害関係者にコンタクトすることが難しい、③利害関係者に研究のリスクや有益性を説明するのが難しい
- 上記のため、個別にインフォームドコンセントを得るのが不可能であると判断する場合、インフォームドコンセントを得ずに研究を行うために、研究者は、利害関係者へのリスクと、どうしてインフォームドコンセントを得るのが不可能であるかを明確に説明できる必要がある。また、研究者は研究によるリスクが最小であること、事前にインフォームドコンセントを得ない代わりに利害関係者をどのように保護するのかを示すと共に、研究により影響を受けた利害関係者に対して彼らの研究への関わりについて適切な情報を事後的に提供する準備が必要である。

ICT研究における「恩恵」の扱い

- 生物医科学に比べて、ICT研究は与える影響の規模が大きくスピードも速く成り得るため、十分なアセスメントが必要である
- **研究が潜在的に与えうる危害の特定：**守秘性(例：個人情報の暴露)、完全性(例：重要情報の喪失)、可用性(例：重要システムの不稼働)の観点で分析する
- **研究が潜在的にもたらす恩恵の特定：**どの利害関係者にとってどのような恩恵があるのか、恩恵は長期的なものか、短期的なものか、という観点で分析する
- **危害の最小化と恩恵の最大化：**適切な匿名化、データ収集の正当性と必要性の確認、収集データの適切な管理、脆弱性情報のkey stakeholdersへの事前通知による修正効果の最大化と悪用リスクの最小化 (Responsible Disclosure)、実システムでの実験の必要性(疑似環境で同様の結果が得られないか)の検証

ICT研究における「正義」の扱い

- 「正義」については、生物医学に比べてICT研究の場合が必ずしも問題が大きいわけではない。
- 研究者は研究の目的に関係のない属性(人種、性別、宗教等)に基づき研究対象を選ぶべきではないが、ICTにおける不透明性(被験者が対面でないこと)は偏った研究対象の選択を防ぐことにつながる可能性がある
- 研究の恩恵がどの利害関係者(攻撃者などは除く)にもできるだけ平等に分配されるようになっているかを考慮する．責任ある開示(Responsible disclosure)は「正義」の観点でも重要．

「法と公益の尊重」について

- 法の順守：ICT研究は国内外の法律、規則、組織ポリシーの対象となる可能性があるが、その適用範囲は明確でない場合も多い。かといって、全ての関連法に基づき研究プロジェクトの合法性を確認するのは現実的でない。
- 合法的な研究を行うために何をすべきか、を検討するよりも、研究により法的に守られた利益を侵害されうるのは誰かを特定すべきである。(=利害関係者の特定)

「法と公益の尊重」について

- 研究がオープンで再現可能であることと、脆弱性の影響を受ける人々を守るために詳細を隠すことをできる限り両立するために適切な脆弱性の開示方法を検討する
- 脆弱性に対して責任がある組織(例：脆弱機器のベンダ)が過去に研究者からの指摘に対してどのように振舞っていたか、すぐに開示しない場合の被害の度合いはどの程度か、脆弱性を修正する前に攻撃が悪用される恐れがどの程度あるか、などを検証する
- **Coordinated disclosure:** 影響を受けうる組織やこれらの組織を保護できる可能性のある組織に対して詳細を開示する
- **Full disclosure:** 脆弱性を完全に一般開示する
- **Closed discussion:** CERT, クローズドな議論を行う

仮想事例で考える

- メンロレポートの補足文書[2]には実際の事例に基づく仮想事例が用意されている。

- 以降では、これに基づき、倫理的に正当性を主張できる(Ethically defendableな)研究について考察する。

- なお、これらの仮想事例のもとになった実際の研究の多くはトップクラスのセキュリティ国際会議で採録されたものである。

仮想事例１：前提となる背景

- (米国の)大学の研究者がボットネットの振る舞いについて総合的な研究を実施している。

- この研究者の目標は、技術的、経済的、社会的な観点でボットネットの拡大、制御、利用を理解することである。

仮想事例１(解析環境検知のための外部攻撃の一部許可)

- 研究者はボットの振る舞いを観測可能なインターネット接続された実験環境(＝ハニーポット/サンドボックス)を作成
- これまでの研究経験から、攻撃者は解析環境であるかを確かめるためにメール送付や他ホストへの攻撃・感染など**悪意のある活動をテストとして実施**することがわかっている。そこで、研究者はこれらの動作の一部を許可して観測を実施する

- **重要な利害関係者：**実験環境から流出する攻撃の被害者(インフォームドコンセントを得るのは無理)
- **恩恵：**ボットネットの動作の詳細観測により感染防止や対策に資する
- **起こりうる危害：**実験環境外に迷惑メールが流出したり、外部へマルウェア感染が発生する

対応する実例

- J. P. John, A. Moshchuk, S. D. Gribble, and A. Krishnamurthy. Studying Spamming Botnets Using Botlab. in Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation (NSDI '09), Apr. 2009

- Stormボットネット(スパムボットネット)の観測のための環境を用意するものの、第三者へのメール送付を完全に防ぐことはできないと判断し、わずかにいくつかの既知のC&Cサーバの接続のみを許可して他の全ての外部向け通信を遮断
- 文献[2]での評価：“By taking a very conservative stance, they are minimizing potential harm yet simultaneously limiting their future ability to do beneficial research.”

仮想事例2(ボットに収集されたファイルの収集)

- 研究者は、観測対象のボットネットは感染ホストから個人情報を含むファイルや著作権を有するファイルが収集され、攻撃者が用意した**ドロップゾーン**に送られていることに気づく

- 研究者は、解析によりこのドロップゾーンにアクセスするためのアクセス先情報、認証情報を得る

- 研究者は、得られたアクセス情報によりドロップゾーンにアクセスし、ボットネットが収集した**全てのファイルを収集**し、詳細に分析する

仮想事例2(ボットに収集されたファイルの収集)

- **重要な利害関係者：**ボット被害者(インフォームドコンセントは無理)
- **恩恵：**悪性ボットによる重要情報収集がどのような規模や内容で発生しているのか実態を詳細に把握でき、対策につなげることができる。実際に漏えいした情報を認識できる。
- **起こりうる危害：**同意がないにも関わらず漏えいしたファイルを研究者に取得されてしまう。これには個人情報や著作権を有する内容も含まれる

対応する事例

- T. Holz, M. Engelberth, and F. C. Freiling. Learning more about the underground economy: A casestudy of keyloggers and dropzones. In M. Backes and P. Ning, editors, Computer Security - ESORICS 2009, 14th European Symposium on Research in Computer Security, 2009.

- 著者らはキーロガーにより集められた情報のドロップゾーンの位置情報を検体内から抽出した。7か月におよぶ実験の結果、ドロップゾーンから17万の感染ＰＣより集められた33GBのデータを収集した。この中には1万の銀行口座と15万のメールアドレスとパスワードが含まれていた。
- 収集された情報はAuSCERTに提出され、被害者には通知が出された。(この手続きをしなければ十分な「恩恵」とは判断されなかった？)

参考文献

1. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research, https://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803.pdf
2. Applying Ethical Principles to Information and Communication Technology Research A Companion to the Menlo Report, http://www.caida.org/publications/papers/2013/menlo_report_companion_actual_formatted/menlo_report_companion_actual_formatted.pdf

Responsible disclosureの具体的な事例紹介(横浜国大の例)

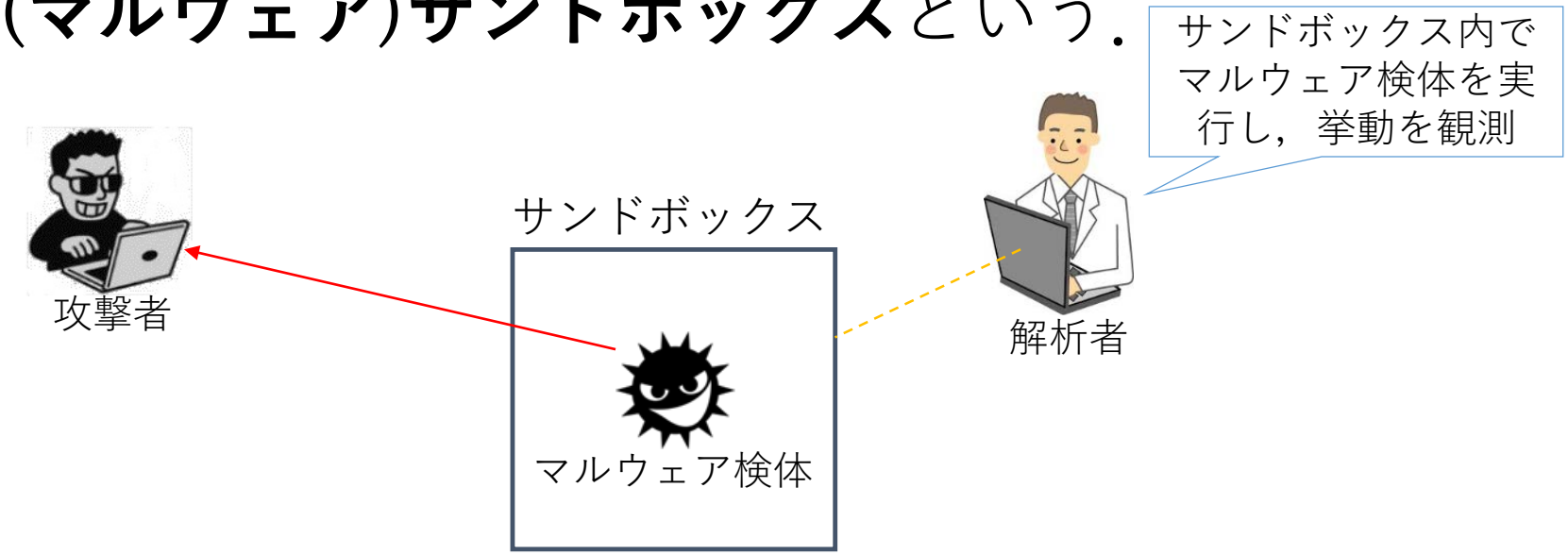
SandPrint: Fingerprinting Malware Sandboxes to Provide Intelligence for Sandbox Evasion

A. Yokoyama, K. Ishii, R. Tanabe, Y. M. P. Pa,
T. Kasama, K. Yoshioka, T. Matsumoto, D. Inoue,
C. Rossow, and M. Backes,

The 19th International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2016), 2016.

(マルウェア)サンドボックスとは？

- **マルウェア動的解析**とは，解析対象のマルウェア検体を解析環境内で実行し，その挙動を観測する解析手法.
- マルウェア動的解析に用いられる解析環境を**(マルウェア)サンドボックス**という.



こんなところで使われています(1)
セキュリティアプライアンス

- ネットワークトラヒックやメール添付ファイルを解析してマルウェアを検知する製品
- **サンドボックス**が内蔵されており、この中で検査対象を実行して悪性なファイルを検知

製造社名	アプライアンス名 / サービス名	種類	製造社名	アプライアンス名 / サービス名	種類
Bluecoat	Malware Analysis System[2]	オンプレミス型	Lastline	Lastline on-Premise[12]	オンプレミス型
Check Point	Threat Emulation[3]	オンプレミス型 / クラウド型	McAfee	Advanced Threat Defence[13]	オンプレミス型
Cisco	Advanced Malware Protection[4]	クラウド型	Paloalto	WildFire[14]	クラウド型
Dell	SonicWALL Capture[5]	クラウド型	Proofpoint	Targeted Attack Protection[15]	クラウド型
FFRI	FFR Yarai Analyzer[6]	オンプレミス型	Secure Brain	Zero-Hour Response[16]	オンプレミス型
FireEye	Malware Analysis[7]	オンプレミス型	Sophos	Sandstorm[17]	クラウド型
Fortinet	FortiCloud[8]	クラウド型	Symantec	Advanced Threat Protection[18]	クラウド型
Fortinet	FortiSandbox[9]	オンプレミス型	TrendMicro	Cloud App Security[19]	クラウド型
Hitachi	MAAS[10]	オンプレミス型 / クラウド型	TrendMicro	Deep Discovery Analyzer[19]	オンプレミス型
IIJ	SecureMX[11]	クラウド型	WatchGuard	APT Blocker[20]	クラウド型
Lastline	Lastline Cloud[12]	クラウド型	Websense	Sandbox Modules[21]	オンプレミス型

田辺瑠偉, 石井攻, 横山日明, 吉岡克成, 松本勉, “標的組織の内部情報を有する攻撃者を前提としたセキュリティアプライアンス評価,” 情報処理学会コンピュータセキュリティシンポジウム2016, セッション3F4, 2016 より

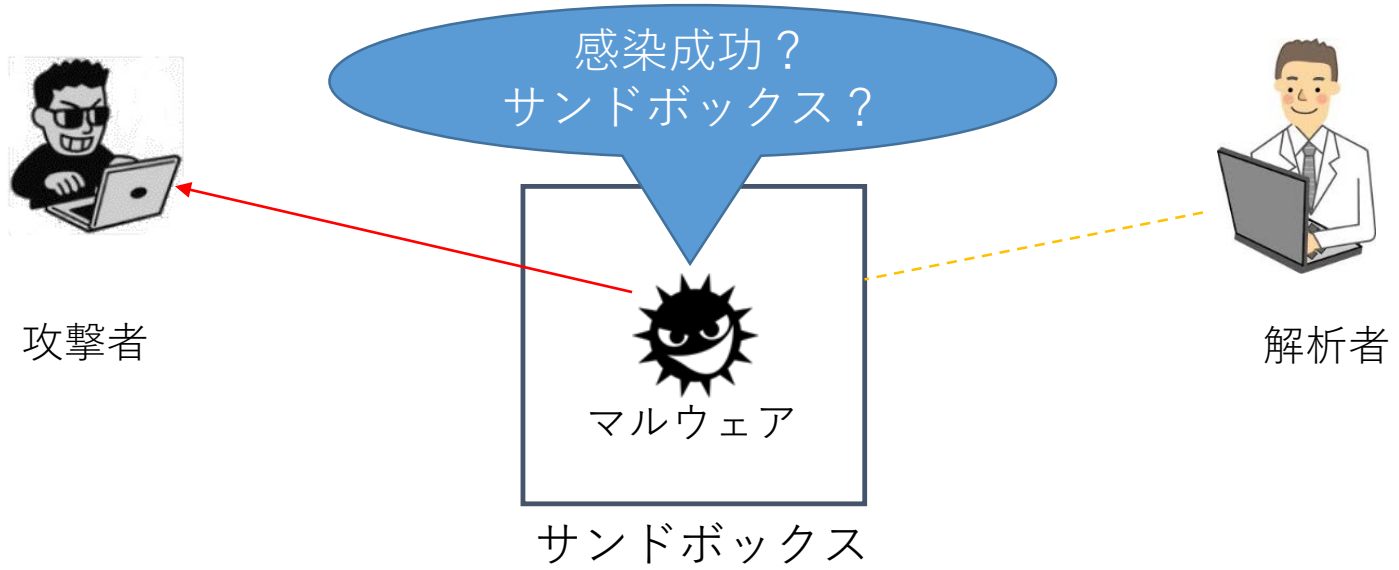
こんなところで使われています(2)

オンラインファイル解析サービス

- ファイルをWebから投稿するとマルウェア解析結果を返してくれるサービス
- バックエンドに**サンドボックス**が動作し投稿ファイルが自動的に解析される
- 有名なサービス(VirusTotal)は一日100万件近いファイル投稿がある

攻撃側と防御側の戦い

- 攻撃側はサンドボックスを検知して無害な振りをするマルウェアを使用して検知や解析を逃れようとする.
- 防御側はサンドボックスであることが見破られないように工夫する.
- これまで攻撃側、防御側の立場で多くの研究



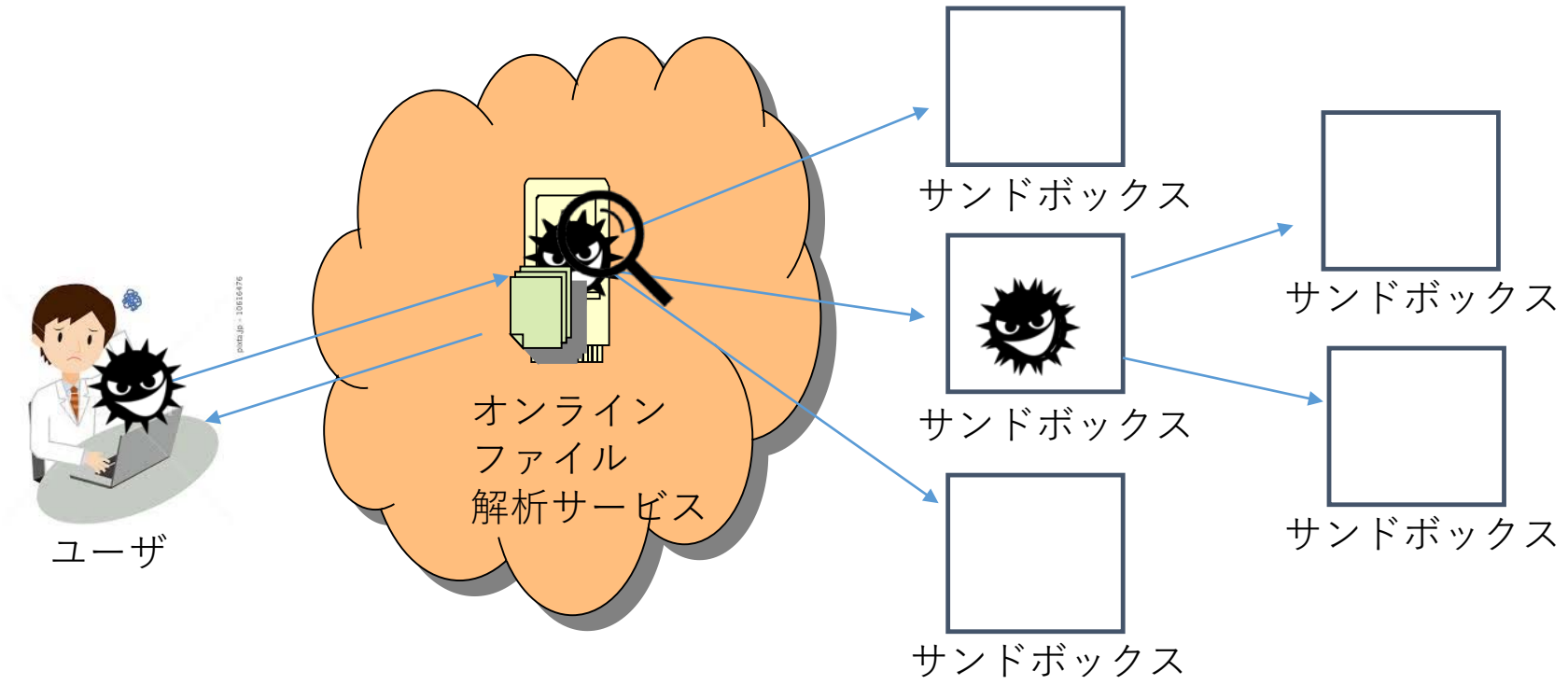
我々が着目したサンドボックスの特徴

サンドボックスは様々な理由で一般ユーザの環境とは異なる特徴を持ち，これらによって解析が検知される恐れがある

- ハードウェア
  - サンドボックスによっては，割り当てられるメモリ等のリソースが限られている.
- スナップショット
  - サンドボックスはスナップショットで状態をマルウェア感染前に復元するため，ファイルアクセス履歴などのユーザの操作履歴が乏しい.
- 環境構成
  - サンドボックスにはデスクトップやOSの設定が初期状態でいわゆる一般ユーザらしさが無い可能性がある.
- ユーザ操作
  - 実行中のマシンの操作が少ないあるいは全く無い.

サンドボックス情報収集ツールSandPrint

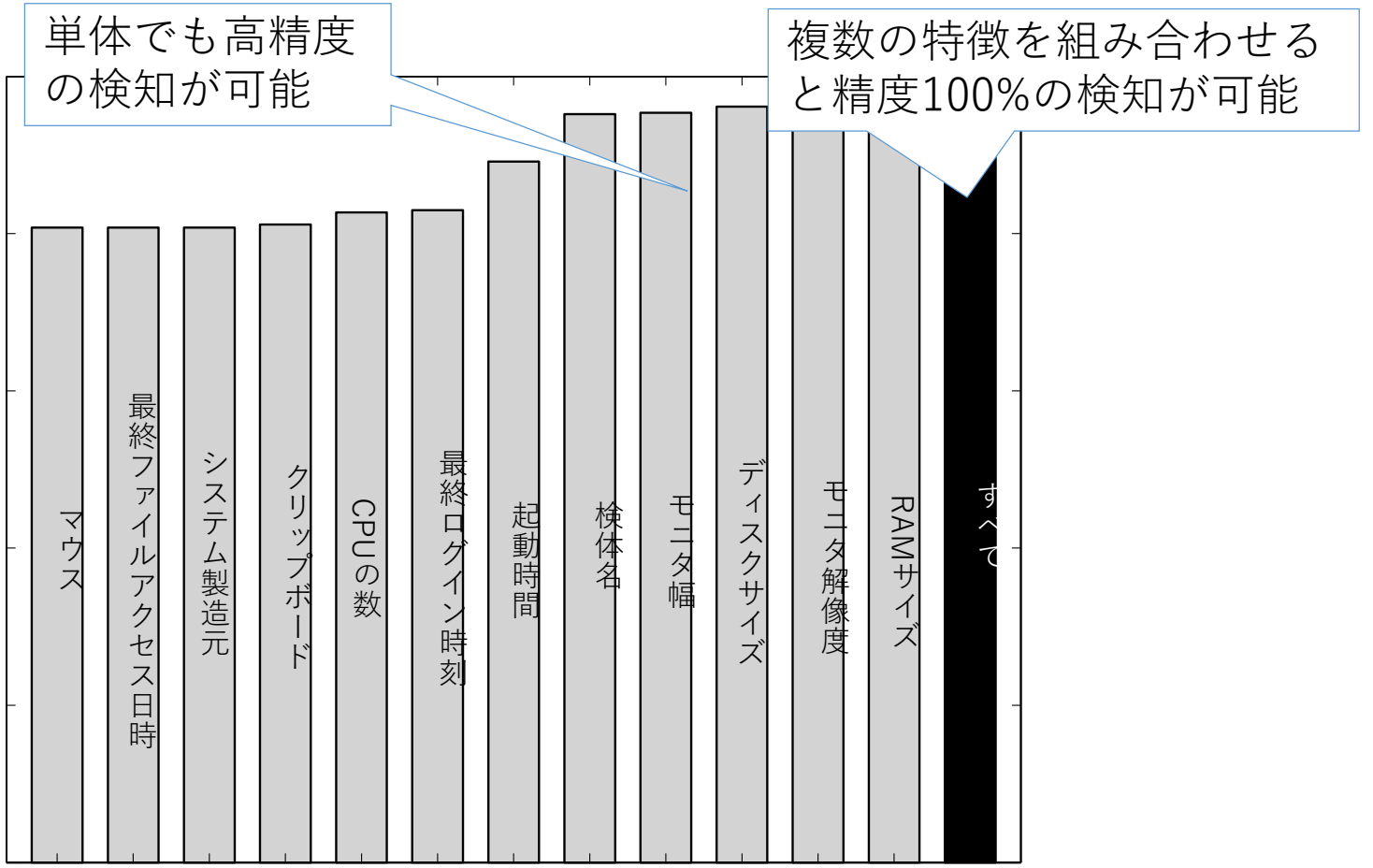
- サンドボックス固有の特徴を調査するために
  サンドボックス情報収集ツール SandPrint を作成.
- オンラインファイル解析サービスに投稿.



結果

- SandPrintにより，合計で2666件の実行環境内の情報に関するレポートを収集した.
- 20サービス中11のマルウェア検査・解析サービスに投稿したSandPrintからの通信を確認.
- 合計 33 カ国の 395 IP アドレスから通信を観測.
- サービスに投稿された SandPrint が共有され,様々な研究機関などで動的解析されていることが確認できた.

高精度でサンドボックス検知が可能



サンドボックス固有な特徴により高精度なサンドボックス検知が可能.

セキュリティアプライアンスの検知

実際に販売されている3つのセキュリティアプライアンスで同様に高精度のサンドボックス検知が可能だった

→サンドボックス構築時にはこれらの特徴に着目した検知に注意する必要があり、セキュリティベンダはこのようなサンドボックス検知への対策を講じる必要があるといえる．

RAID2016投稿の裏話

- 論文投稿時は製品名・サービス名は全て実名で投稿(ただし、出版時には製品名を匿名化する予定であることを記載)

- プログラム委員会から「この対応では不十分」との指摘．適切な「Responsible Disclosure」対応をしなければ採録とできない旨のコメントを受ける

- シェパード(論文添削監視者：お世話役)は、イリノイ大学のMichael Bailey准教授であり、**メンロレポートの著者の一人**だった

Responsible Disclosureと論文採録までの流れ

2016/6/4　論文が条件付き採録となりResponsible Disclosureが採録要件となる

2016/6/10　シェパードと著者の各組織(横浜国大、NICT、ザールラント大)が一同にSkypeミーティング。今後のDisclosureの手順について提案し、シェパードから承認を得る

2016/6/10-20　Disclosureの準備(研究内容と指摘する問題点の説明文、連絡先情報等の確認)

2016/6/21 Disclosureの開始(セキュリティベンダ３社、オンライン解析サービスオペレータ２０組織が対象)

Responsible Disclosureと論文採録までの流れ

2016/6/29 シェパードへの中間報告(この時点でセキュリティベンダ３社全て、解析サービスオペレータ20社のうち11社から返信あり)

2016/7/4　返信の無い解析サービスオペレータ9社に他のチャネル(組織の問い合わせ窓口メールなど)から再度連絡

2016/7/9シェパードとプログラム委員会により採録の判定を受ける

2016/7/18 返信の無い解析サービスオペレータがある国のNational CERTに連絡

2016/9/19-21 RAID2016開催(論文の公表)

スケジュール

90日間

３つのセキュリティベンダと２０のオンライン解析サービスのオペレータに研究内容と脆弱性について通知(セキュリティ情報提供用メール、または、Webフォームより)

14日間返答がない場合

別のチャネル(各企業の問い合わせ用アドレス等)で連絡

14日間返答がない場合

通知先組織がある国のNational CERTにメール連絡

RAID 2016

論文内での記述(約1ページ)

7.1 Ethical Considerations

Our research may seem offensive in the sense that we reveal fingerprints of malware sandboxes that adversaries can use to evade them. Note, however, that the information we presented can be gathered by any other person reproducing our (conceptually simple) fingerprinting method. We thus consider the information shown in this paper as public knowledge. Still, we present data only in aggregated form and refrain from revealing any internals of particular sandboxes.

Using our insights, sandbox operators can
ysis systems. For example, we have shown th
features that are inherent to the snapshot o
be possible to find artifacts that can identif
nificantly harder to build a classifier that wo
more people randomize characteristics. We h
ticularly characteristic of sandboxes, giving
to significantly improve the stealthiness of th

7.2 Responsible Disclosure

Organizations developing sandboxes and/or a
by our research results and we thus consider
responsible disclosure process. To notify these
90 days prior to the publishing date of this pa
and including hints on how to protect against potential adversaries in the future. We used direct contacts whenever possible and available. Alternatively, we resorted to contact details stated on the organization's websites, notably including Web-based contact forms. If we did not receive a response after 2 weeks, we retried to contact the organization, if possible using alternative communication channels (e.g., using generic email addresses like `info@organization.com` or email addresses found in the WHOIS database for the organization's website domain). If we did not hear back from the organization after 4 weeks, we contacted the national CERT(s) that are in the same country as the affected organization in order to notify the party via the CERT as trusted intermediary.

We handed to each organization an executive summary of our research results as well as a full description of our research methodology (i.e., a copy of this paper in the pre-print version). We made sure to highlight the implications of our work with respect to future operations of the sandbox and/or appliance. We also specified our contact details of both research institutions, including physical address, phone number, and the email address of a representative for the research activities. We allowed the organizations to download the latest version of SandPrint and its source code. Such auxiliary data is helpful to build protection mechanisms against sandbox-evasive programs similar to SandPrint. We also remove all organizations' names when referring to individual sandboxes/services.

開示先組織からの反応

- **最終的に１８の組織から研究に対する**
  ポジティブな意見と反応が得られた。

- ネガティブな反応はなかった。

- SandPrintのソースコードを７つの組織に提供した。

- その後、我々が提案した検知に対応した製品が複数確認できた。

脆弱性開示を伴う論文投稿の経験から学んだこと

- Responsible Disclosure等の適切に対応すれば、脆弱性指摘に対するベンダの反応はポジティブなものが多かった(友好的な反応を示した組織のうちの１社は以前に匿名化なしに同社製品の脆弱性を詳細暴露した研究者を訴えた実績があった)

- Responsible Disclosureには手間と時間が掛かるので、脆弱性研究を発表する場合は、**時間に余裕をもって対応を計画**する必要がある。ただし、論文が採録されないうちに開示を始めると、論文がリジェクトされた際に脆弱性情報だけ世の中に伝わってしまい、研究の新規性が失われる可能性もある。したがって、投稿時には脆弱性開示の方針を論文に示して、採録された際に速やかに開示を始めるという選択肢もある。(世の中を安全にするという目的だけを考えれば、投稿前開示でも良いが、世界の研究者は競争の激しい国際会議で成果を発表し続けなければ評価されないというプレッシャーがあり、このあたりはぎりぎりの選択と言える。)

- 必要以上に倫理問題を意識して保守的な研究を行うよりも、世界の動向は**必要な責任を果たして社会への恩恵を高める**研究を評価する傾向にあるのではないか。(＝メンロレポートの考え方)

事例２：DDNS脆弱性

M. Korczynski, M. Krol, M. van Eeten, “Zone Poisoning: The How and Where of Non-Secure DNS Dynamic Updates,” ACM Internet Measurement Conference, IMC2016, 2016.

謝辞：本事例は著者の一人であるデルフト工科大学Michel van Eeten教授から伺った事例です。本講演で上記論文の査読プロセスにおけるやり取りの一部に触れることについては、ご本人の了解を取っています。ご快諾いただいたvan Eeten教授への謝意を表します。

DNS Dynamic Updateの脆弱性

- RFC2136(1997) レコードの更新は更新パケット(UDP!)の送信元IPアドレス等の条件を満たせばだれでも可能．無条件に更新を受け付ける設定も可能.
- RFC2535(1999) DNSSECによる認証
- RFC2845(2000) HMACによる認証(TSIG)

- BIND: v8とv9ではデフォルトでオフだがany設定等で任意のIPアドレスからの更新を受け入れる設定も可能, v9.1からスレイブからマスタ権威サーバへの更新パケット転送に対応
- Windows Server: 2000からDynamic Updateを導入．デフォルトではTSIGによる更新が有効だが、認証なしにも設定可能.

ＤＮＳレコードの更新(追加含む)が自由にできれば、いうまでもなく、様々な攻撃に悪用可能=Zone Poisoning攻撃

脆弱性(設定ミス)調査方法

- 世界中で運用中の権威DNSサーバの設定を調査し、この問題(Zone Poisoning脆弱性)の実情を調べたい
- 脆弱性が存在するか調べる唯一の方法は実際にDynamic Updateを試してみるしかない
- 管理者の許可を得て行う調査ではスケールしない．世界中の権威ＤＮＳサーバにテスト用のレコード追加を(管理者の許可を得ずに)実際に行い、脆弱性の現状を調べるしかない！

実際にやりました。

調査の結果

- DNSDB等からランダム抽出した290万のドメインとAlexaトップ100万ドメインを調査し、それぞれ1,877件(0.065%)、587件(0.062%)がZone Poisoningに脆弱であることがわかった。

- これらのドメインの中には政府系、医療機関、銀行など、攻撃により深刻な影響が想定されるものが含まれていた。

どうやって、このような調査を
”倫理的”に行ったのか？
（＝主要カンファレンスで受け入れられたのか？）

著者らが行った倫理的研究のための対応 (論文に書かれていること)

4.3節 Ethical Considerations (全文)

While vulnerability scanning has become an established part of security research, our approach does raise ethical questions because of the fact that the only valid method available to us for assessing the vulnerability of a DNS server was to add a record to the zone file.

脆弱性スキャンは確立されたセキュリティ研究の一部であるが、我々のアプローチは倫理的な疑問を生じさせる．ＤＮＳサーバの脆弱性のアセスメントには実際にゾーンファイルにレコードを追加してみる必要があるからである．

We have submitted the study to the TU Delft Human Research Ethics Committee. The committee evaluated our request and stated that we did not need their authorization since we were not conducting human subjects research.

我々はこの研究内容を(著者が所属する)デルフト工科大学の倫理委員会に提出したが、委員会の判断は、「人間を対象とした研究ではないため、当委員会の承認を得る必要はない」という判断だった。

著者らが行った倫理的研究のための対応 (論文に書かれていること)

While this makes sense, it also signals that current institutional review procedures are not set up to evaluate ethical issues in computer security. We have assessed our work using the principles outlined in the Menlo report.

この委員会の判断は理にかなっているが、コンピュータセキュリティの研究において、既存の倫理委員会が適切に機能を果たしていないことを懸念させる。そこで、我々は、メンロレポートにある研究倫理を用いて、自らの研究を評価することとした。

We do not collect data on persons. Getting informed consent before adding a record to the zone file is both unpractical and would introduce selection bias, since administrators of well secured servers are more likely to consent.

我々は個人のデータの収集はしていない。インフォームドコンセントを実験の前に得るのは現実的でないし、事前に同意を得る実験では、結果にバイアスが掛かる懸念がある(調査を受け入れる管理者はセキュリティ対策を行っている可能性が高いため)

著者らが行った倫理的研究のための対応 (論文に書かれていること)

We do provide a clear opt-out mechanism via the website referenced in the added DNS record. The site also provides full transparency regarding the study and its objectives.

我々は、研究内容と目的を包み隠さずに説明した研究説明用Ｗｅｂサイトを用意し、(検査対象のＤＮＳサーバの管理者が閲覧できるように)脆弱性検査のために挿入したレコードにこのサイトの情報を記述した。また、このサイトには、我々への連絡先情報を記載し、当該調査を今後拒否するという管理者の希望があれば、これを受け入れた(オプトアウトメカニズムの導入).

Our approach in testing the vulnerability has been designed to have as minimal impact as possible: we send a single RFC-compliant packet. We do not read, change or otherwise engage with any existing records.

我々の脆弱性検査のアプローチでは、検査対象に対するインパクトが出来るだけ小さくなるようにしている。各サーバに対してRFCに準拠した、ただ１つのパケットを送るだけであり、情報を読み取ったり、変更したり、既存のレコードに対してどのような関与もしていない。

著者らが行った倫理的研究のための対応 (論文に書かれていること)

We feel the drawback of lacking consent from server operators is outweighed by the benefits of our measurement for those operators: to be made aware of a critical vulnerability in their DNS server.

事前にDNS管理者に同意を得ることができないという問題よりも、深刻な脆弱性について彼らに情報を提供できるという恩恵の方が大きいと考える

All notifications have been completed before the publication of this paper. The new record is highly unlikely to be discovered by accident and it is removed at the end of the study.

論文の出版の前に、脆弱性を有する権威サーバの管理者に対する通知は全て完了した．実験のために挿入されたレコードが偶然に発見され、運用に支障をきたす可能性は非常に低く、実験後にそれらのレコードは削除された(ことを確認した).

プログラム委員会，査読者の反応

Review #A

査読結果は配布資料から割愛させて頂きます。

この事例からわかる事

- Internet Measurement分野のトップ会議であるACM IMCのPCメンバですら、倫理的な研究に関する評価は**査読者間で１８０度異なる**場合がある(査読者Dは研究の違法性を否定的に指摘したが、他の査読者はむしろ倫理的対応を適切とし、本研究の強みとしている)

- 大学の既設の研究倫理委員会(IRB)が、サイバーセキュリティ研究の倫理に対応していないのは、日本だけではない。また、ドイツの共同研究者に同様の質問をしたところ、オランダに限らず、ドイツも同じであり、サイバーセキュリティ研究倫理を評価できる体制が整っていないとのこと。→**「大学のIRBに承認を得たため倫理的に問題ない」という説明は、通用しない場合がある**。

この事例からわかる事

- 査読者Ｂの「組織のIRBによる研究倫理評価の限界ではなく、**著者自身の研究(の意味すること)への理解や評価が不十分なのではないか**」というコメントが最も重要と考える

- 「研究倫理の考察・対応」は研究のメインの活動ではないので、できれば簡単に済ませたいのが研究者の本音
- **自分では何も考えずに「ガイドライン」に従って安全に研究したい！　誰か「ガイドライン」を作って！と他力本願になりがち。**
- しかし、このような考え方は**無責任**であり、社会的に受け入れられないだけでなく、競争の激しい難関会議では通用しない。
- サイバーセキュリティ研究は千差万別であり、ガイドラインで個々の研究を評価しお墨付きを与えるのは無理. **自分の研究の意義やインパクト、問題点は、自分自身で最も深く、正しく、厳しく、検討・評価し、それを世に問い続けることで、自身の研究倫理に関するセンスを磨くしかない(自分自身への戒め)**

システムセキュリティ系トップ会議では

- ACM CCS 2017 CFPより

If a paper includes work that raises ethical concerns **it is up to the authors to convince the reviewers** that appropriate practices were followed to minimize possible harm and that any harm caused by the work is **greatly outweighed by its benefits.**

システムセキュリティ系トップ会議では

- Usenix Security 2016 CFP(IEEE S&P2017 CFPでも参照)より
- Human Subjects and Ethical Considerations Submissions that describe experiments on human subjects, that analyze data derived from human subjects (even anonymized data), or that otherwise may put humans at risk should
  - 1. Disclose whether the research received an approval or waiver from each of the authors' institutional ethics review boards (IRB)— if applicable.
  - 2. Discuss steps taken to ensure that participants and others who might have been affected by an experiment were treated ethically and with respect.
- **If the submission deals with vulnerabilities (e.g., software vulnerabilities in a given program or design weaknesses in a hardware system), the authors need to discuss in detail the steps they plan to take to address these vulnerabilities (e.g., by disclosing vulnerabilities to the vendors).** The same applies if the submission deals with personal identifiable information (PII) or other kinds of sensitive data. If a paper raises significant ethical and legal concerns, it might be rejected based on these concerns.

システムセキュリティ系トップ会議では

- Usenix Security 2016 CFP(IEEE S&P2017 CFPでも参照)より
- Authors seeking ways to reduce the ethical risks of their experiments may **optionally consider reaching out to the Ethics Feedback Panel** for Networking and Security at www.ethicalresearch.org/efp/netsec/. The panel's mission is to help researchers identify ethics-related risks, find prior research that provides precedent or data to inform ethical decision making, to suggest ways to improve experimental designs to reduce ethical risks, and provide any other information that may assist the researchers in meeting their ethical obligations.
- The best time to reach out to this panel is **before conducting your experiments, but they may be able to assist if concerns arise during an experiment.** Contact the program co-chairs at sec16chairs@usenix.org if you have any questions

CSS2018 (情報処理学会コンピュータセキュリティシンポジウム2018)でも、、

- **研究倫理相談窓口**の設置を検討中

- 研究倫理について、どのような点で研究者が困っており、どのように問題を評価すればよいかを一緒に議論し、検討する場として期待

- 設置自体の是非や、実際の運営の仕方など、皆さまのご意見を頂けますと幸いです。

- ご協力頂ける方も募集中です。