サイバーセキュリティ研究倫理と日本における活動

NTT セキュアプラットフォーム研究所
特別研究員
秋山（あきやま）　満昭（みつあき）

横浜国立大学　大学院環境情報研究院
准教授
吉岡（よしおか）　克成（かつなり）

1. サイバーセキュリティの研究倫理

インターネットの急速な発展は、私たちの生活環境を便利で効率化した反面、サイバー攻撃の増加や個人を特定できる情報の増加をもたらしている。また、ICT研究、特にサイバーセキュリティ研究に関しては、研究過程やその結果が私たちの生活環境に重大な影響を与えるものとなっている。

サイバーセキュリティの研究を行うためには、インターネット上の事象を観測することが欠かせない。しかし、このような観測研究においては、攻撃通信だけでなく様々な通信が含まれ得るため、プライバシーをはじめとした様々なセキュリティリスクが存在する。このためリスクを最小化するための実験設計や、リスクを被る可能性のある対象に十分な説明責任を果たした上で実施する必要がある。

また、研究の過程で特定のソフトウェアのセキュリティホール（セキュリティ上の欠陥）を発見した場合、影響を受けるシステムやサービスに対して十分な対処ができていない状態で明るみに出た場合には、攻撃に悪用されるリスクがある。

インターネットの普及に伴い、2000年代初頭から大規模なICT研究が実施されてきた。その一方で、研究倫理に関する適切な指標がないため、倫理的検証が不十分な研究があったことも否めない。このような状況に鑑みて、2010年代からはICTやサイバーセキュリティに関する研究倫理の議論が米国を中心に盛り上がりを見せており、新たな倫理原則が策定されたことや、学術研究論文に対して研究倫理の観点からのレビューなどが行われている。サイバーセキュリティにおける革新的な研究は研究倫理と表裏一体であり、世界で競争力のある技術を創出するためには、研究倫理は研究者や技術者にとっての必須の素養になっている。

また、近年では日本においてもサイバーセキュリティの研究倫理に関する重要性が広く認識されつつある。学術機関を中心としたシンポジウムの実施や、サイバーセキュリティの研究倫理を啓発し推進する組織が設立されている。

本稿では、サイバーセキュリティ研究倫理の原則について解説し、世の中の動向や最新事例を紹介し、日本におけるサイバーセキュリティ研究倫理の状況とこれからの展望について述べる。

2. 研究倫理原則

初期のICT研究では、研究倫理に関する適切な指標がなく、倫理的検証が不十分なまま行われたものも多く存在する。例えば、マルウェアの取扱い、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。このような状況に鑑みて、1979年に制定された生物医学分野の研究倫理を定めた「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た。そして米国研究者を中心としてまとめられ、2012年にアメリカ合衆国国土安全保障省から「メンロレポート」が発行された。本章では、ベルモントレポートの研究倫理原則を紹介し、生物医化学とICT研究の違いを考察した後に、メンロレポートで新たに規定された研究倫理原則を紹介する。

2.1　ベルモントレポート

ベルモントレポートは生物医学分野の研究倫理を定めたもので、研究倫理原則として、以下のように示されている。

・人格の尊重（Respect for Persons）
研究への参加は本人の自由意志によって決まり、本人への十分な説明をした上で本人が意思決定する権利を尊重するというインフォームドコンセントの考え方。

・恩恵（Beneficence）
研究により得られ得る恩恵を最大にし、与え得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る。また、研究の恩恵は平等に分配され、負担は研究対象に対して同等に分担されるべきである。

2.2　生物医化学とICT研究の違い

ベルモントレポートで言及されている倫理原則は、生物医科学研究をベースとするものの、他分野でも広く適用できる基礎的な行動規範を示唆している。しかしながら、ICT研究は、ベルモントレポートが制定された当時では想定し得なかった環境条件に基づいて実施されることに注意しなければならない。具体的には、生物医科学とICT研究には以下に列挙する違いがある。

・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、対象は数十から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る。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容易ではない。

・速度

生物医学では多くの研究はマニュアルプロセス（研究室において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研究では瞬時に数百万規模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の迅速かつ的確な判断を要する。

・情報の集約と相互関係

ICT研究において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、メールアドレス、友人等の連絡先リスト、SNSアカウント等の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つながる可能性がある。

・分散化

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。

・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。

このようなICT研究における状況の違いを踏まえて、研究者は倫理的に研究計画を設計し実施しなければならない。

2.3　メンロレポート

インターネットの発展に伴い、生物医化学とICT研究の違い（2.2節）が鮮明になるにしたがって、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た。そこで、ベルモントレポートを踏襲しつつ、ICT研究に対応したメンロレポート[1]が2012年に制定された。メンロレポートでは、3つの原則をICT研究の文脈で解釈することに加えて、下記の研究倫理原則を新たに追加している。

・法と公益の尊重（Respect for Law and Public Interest）

法令を遵守すること。公共の利益を尊重すること。研究方法と結果の透明性を保ち、その行為に責任を持つこと。

この倫理原則は、ベルモントレポートにおける“恩恵（Beneficence）”に関して、ICT研究の文脈で解釈し、新たに取り組むべき課題として、地域間における法律の対立や曖昧さ、利害関係者（ステークホルダ）の特定の難しさ、法と公益の間の不一致、などを明示したものである。研究過程でセキュリティホールを発見した場合は、それの影響を受け得るステークホルダを特定して、被害を最小化するための責任ある情報開示（Responsible disclosure）を実践しなければならない。

なお、メンロレポートの付属資料として、事例に基づいた議論と対応がまとめられている[2]。

3. 世の中の動向

サイバーセキュリティの研究倫理への関心の高まりを受けて、サイバーセキュリティ研究倫理をテーマとした新たな学術国際会議としてCREDS、CREDSII、NS-Ethicsなどが2013年から2015年にかけて開催された[3] [4] [5]。これらの会議では、研究倫理の議論が不十分だった過去の研究事例を再レビューや、ベストプラクティスの共有、倫理的研究の設計、などを実施し、ICT環境の変化に対応した、より良い倫理的研究をサポートしている。

また、サイバーセキュリティのトップ学術会議（IEEE S&P、ACM CCS、USENIX Security、ISOC NDSS）をはじめとする学術会議では、研究倫理に関してCall for paperにおいて言及されることが2013年頃から徐々に増加してきている。

具体的にはCall for paperにおいて、「研究倫理の議論を喚起しそうな論文については、研究倫理に関する明確な記述をすること、また自組織の研究倫理委員会で承認を受けること」を求めている。

では、実際に世の中のサイバーセキュリティ研究ではどのように倫理的に研究を実践しているのだろうか？過去数年

分の難関国際会議論文（USENIX Security 2012 ～ 2016で発表された約300本の論文）において研究倫理に関する記述を調査した結果、研究倫理の議論や主張は大まかに以下のように分類できた。

・同意／承認の獲得
　ユーザ（被験者）の同意、サービス事業者の承認、研究倫理委員会（IRB）の承認
・手順の正当性
　匿名化の実施、ポリシー／ガイドラインの準拠、適法性の主張、代替手段なし、Responsible disclosureの実施
・リスク／被害のコントロール
　リスクの最小化、新たな被害の防止
・利益
　ベストプラクティスの共有、公益性がある
・その他
　人間を対象にしていない、研究用途である

このように先人たちが実践してきた具体的な方法は、これから同種の研究を開始する研究者や技術者にとっての参考にすべき事例集として利用できる。

4. 研究事例紹介

日本の研究組織が行った研究の中でも、特に研究倫理に関わる事例を紹介する。

4.1 サンドボックス検知手法

横浜国立大学の横山氏は、マルウェア解析や検知に用いるサンドボックス（検査対象プログラムを実行してその挙動を調べることでマルウェアを検知したり、その機能を分析するための解析用実行環境）が有する典型的な特徴を明らかにし、これらが悪用されることで、サンドボックスによるマルウェア解析や検知が阻害される恐れを指摘している[6]。

この研究では、まず実際に運用されているサンドボックスの実情を調査するため、任意の検体をWebサイト等で受け付けてサンドボックス解析を行うオンラインマルウェア解析サービスに情報収集用の検体を投稿し、それらのサービスで利用されているサンドボックスの特徴情報を収集している。収集された特徴情報に基づき、機械学習によりサンドボックスと一般ユーザの環境が高い精度で判別されることが示され、さらに、製品化されたサンドボックスに対しても前述の判定器が有効であることが報告されている。

これらの実験結果や実験用検体は、事前にサンドボックス製品ベンダやマルウェア解析サービスプロバイダに提供され、製品やサービスの改善に貢献している。また、論文において、これらの製品やプロバイダの名称、固有の内部情報は匿名化されるとともに、収集された特徴は統計情報として提示され、特定の製品やサービスへの影響を抑えるよう配慮されている。このように研究成果の公表による恩恵をできるだけ増やし、危害を減らす努力がなされている。

4.2 ソーシャルアカウント特定手法

NTTセキュアプラットフォーム研究所の渡邉氏は、ソーシャルWebサービスにおいて、任意の標的ユーザのアカウントを特定する新たなプライバシー攻撃を発見した[7]。この攻撃は、ソーシャルWebサービスに標準的に備わっているブロック機能を悪用するため、広く世の中のソーシャルWebサービスに影響を与える可能性があり、その利用者が攻撃の標的になり得る。

この研究では、攻撃を実証するために実際のサービス上で実験を行っているが、その際に影響やリスクを最小限にするように注意深く実験の設計を行っている。具体的には、実際のユーザに対して攻撃することがないよう自身が用意したアカウントで実験をすることや、不用意にサービスへの負荷が上昇しないような実験設計を行った。

また、影響を受けることを確かめた12サービス事業者に対して事前にコンタクトを行い、攻撃手法の再現方法と対策手法について情報共有を実施した。その後、この攻撃手法を公表することに関してネガティブな意見や延期等を求める事業者はいなかったことや、広く世の中に対して脅威が認識されることによる公益性を考え、国際会議IEEE EuroS&P 2018にて発表された。

5. 日本の現状とこれから

日本では、2016年から国内研究コミュニティである「マルウェア対策研究人材育成ワークショップ（MWS）」[8]において、サイバーセキュリティ研究倫理に関する知見の共有と倫理的研究の実践について本格的に議論され始めた。

その後、国内最大規模のセキュリティ系学術シンポジウムである「暗号と情報セキュリティシンポジウム（SCIS）」[9]や日本学術振興会（JSPS）の公開シンポジウムにおいて、サイバーセキュリティの有識者が一堂に会して、課題の認識や今後のアクションプランが議論された。しかしながら、日本においては研究倫理の議論を要する研究自体がまだまだ少ないため、最初の一歩を踏み出して後押しするための研究的土壌が必要とされている。

5.1　見えてきた課題

・実践と知見共有

世の中へのインパクトや利益は状況によってケースバイケースであり、画一的に判断できることではない。そのため、研究者／技術者間でケーススタディを積み上げる必要がある。特に、Responsible disclosureはステークホルダの見積もりから実施の手順にいたるまで複雑なこともある。日本では、IPAとJPCERT/CCが行う「情報セキュリティ早期警戒パートナーシップ」の窓口に報告することで、製品やソフトウェアの脆弱性に関するResponsible disclosureが円滑に遂行できる仕組みがある。一方で、ソフトウェアや製品の脆弱性とまでは言えないが、Responsible disclosureが必要な場合（4章の事例）もあり、研究者が主体で実施しなければならない状況もある。十分なノウハウ蓄積を単一の組織で実施することは困難であるため、組織横断的に議論をして知見を共有する場が求められている。

・研究的意義

研究倫理の議論が必要な研究においては、単に発見した攻撃手法や脆弱性に関してだけではなく、それを発見した情報工学的手法に研究的価値があるはずである。発見手法を広く世の中に伝えることで、例えば開発者自身がチェックできたり開発段階で発見できたりするなど世の中に対する恩恵がある。また共通的な落とし穴(Common pitfall）とその根本的な対処方法を明らかにすることにも研究的価値があると言える。問題／対策を一般化して世の中に広めることは後世に対して恩恵がある。

・産学連携

学術界と産業界の連携によって初めて実質的な対策ができる。そのためには、産業界を巻き込んだ議論が必要になる。例えば、業界ごとに対策までの猶予期間（Grace period）やResponsible disclosureの方法に関するコンセンサスは異なる可能性がある。また、産学の信頼関係構築にむけて研究者が努力すべきこともある。例えば、Responsible disclosureにおいて十分な情報と猶予期間、ワークアラウンドを提示することなどである。

5.2　これからの取組み

これまでの国内の取組みの重要性が広く認められる中で、JSPSのサイバーセキュリティ第192委員会[10]において、「サイバーセキュリティの研究倫理を考えるWG」というワーキンググループが2018年2月に設置された。このワーキンググループは、JSPSという中立的な組織を母体として中立的な立場から、学術界及び産業界横断的にサイバーセキュリティ研究倫理の理解と実践の促進を支援する活動を推進することを予定している。

サイバーセキュリティの研究倫理に関する判断が正確にできる研究倫理委員会を保有する研究機関は世界的に見ても限られている。このような状況に鑑みて、サイバーセキュリティの研究倫理に関する相談窓口を国内シンポジウムの「コンピュータセキュリティシンポジウム（CSS）」において試験的に設置することが検討されている。研究者は研究を実施する際に、研究倫理に関する不明点がある場合にこの窓口にあらかじめ相談できるものである。

6. おわりに

ICT研究やサイバーセキュリティ研究を進める上で、メンロレポートに規定されている倫理原則は研究者・技術者が持ち合わせるべき必須の知識になっていると言える。

サイバーセキュリティにおける倫理的研究を推進する国内独自の活動が、日本発の先進的・競争力のあるセキュリティ技術の持続的な創出に貢献することを期待したい。

参考文献

[1] Menlo Report, *https://www.caida.org/publications/papers/2012/menlo_report_actual_formatted/*

[2] Applying Ethical Principles to Information and Communication Technology Research, *http://www.caida.org/publications/papers/2013/menlo_report_companion_actual_formatted/menlo_report_companion_actual_formatted.pdf*

[3] Cyber-security Research Ethics Dialog & Strategy Work-shop（CREDS 2013）, *http://www.caida.org/workshops/creds/1305/*

[4] Cyber-security Research Ethics Dialog & Strategy Work-shop（CREDS II-The Sequel）, *http://www.caida.org/workshops/creds/1405/*

[5] Workshop on Ethics in Networked Systems Re-search（NS-Ethics）, *https://conferences.sigcomm.org/sigcomm/2015/netethics.php*

[6] Yokoyama et al., SandPrint：Fingerprinting malware sandboxes to provide intelligence for sandbox evasion, RAID 2016.

[7] T. Watanabe et al., User Blocking Considered Harmful? An Attacker-controllable Side Channel to Identify Social Accounts, IEEE EuroS&P 2018.

[8] マルウェア対策研究人材育成ワークショップ（MWS）, *https://www.iwsec.org/mws/2017/*

[9] 暗号と情報セキュリティシンポジウム（SCIS）, *https://www.iwsec.org/scis/2018/*

[10] 日本学術振興会サイバーセキュリティ第192委員会, *https://www.jsps.go.jp/j-soc/list/192.html*