CSS2018研究倫理窓口における知識ベースの活用(案)

セコム㈱IS研究所
島岡政基

倫理窓口の懸案事項

- 相談内容の整理
  - 投稿前で原稿がない状況での相談
  - 最終的には論文化されるはず(不採録にならない限り)
- 次年度への引き継ぎ？
  - 形式上は年度毎に招集・解散
  - 相談情報は誰が管理してどう引き継ぐか
- セキュリティコミュニティへのフィードバック
  - 窓口メンバ以外にも多少なりとフィードバックができると嬉しい
    - e.g., CSS会期中に企画セッションを設けて報告を行う
  - フィードバック先と開示範囲についても議論が必要
  - 相談者に対する同意の取得や公開方法の模索
  - 初年度から考える話ではないかもしれないが、要念頭

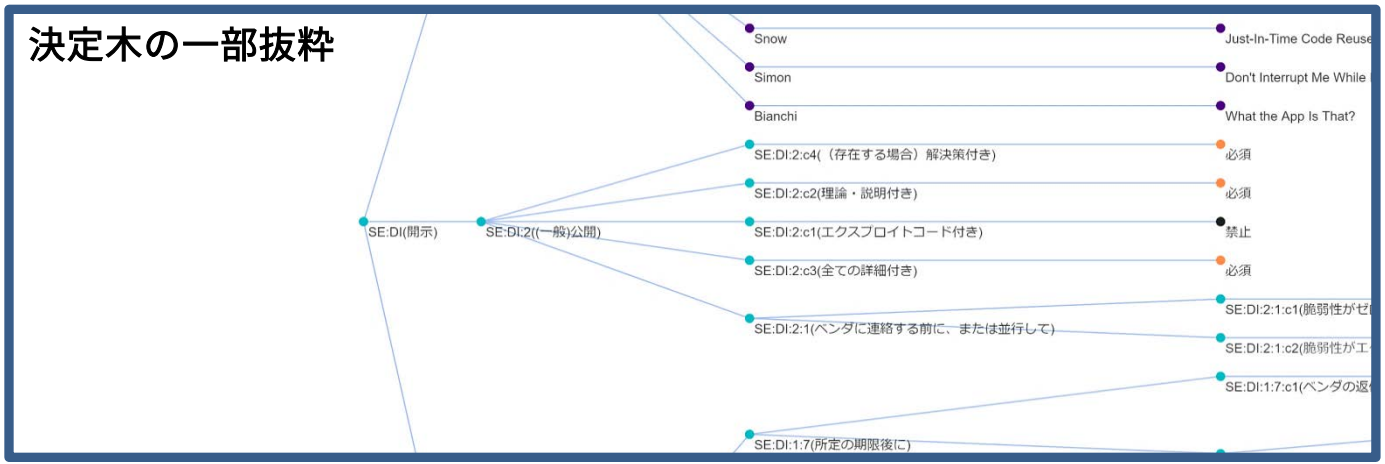
倫理窓口への相談情報を、(できるだけ多くの方々に)
活用できる形で蓄積していきたい
(未整理のまま蓄積していくと後が大変)

相談情報の蓄積方法(案)

1. 生データ(メールベースのテキスト)
   - 相談者と倫理窓口とのメール(css2018-ethic)
   - 倫理窓口メンバの議論のメール(css2018-ethic-wg)
2. 生データを整理した何らかの知識ベース(KB)の構築
   - 生データを次年度以降も効率的に参照できる形にしたい
     - 可能であれば、一定の加工を施した上で窓口メンバ以外も活用できる形にできるとベスト(加工方法、開示範囲など要議論)
   - 継続的な運用可能性を考慮する
     - 蓄積、参照ともに少ない負担で使える方法が望ましい

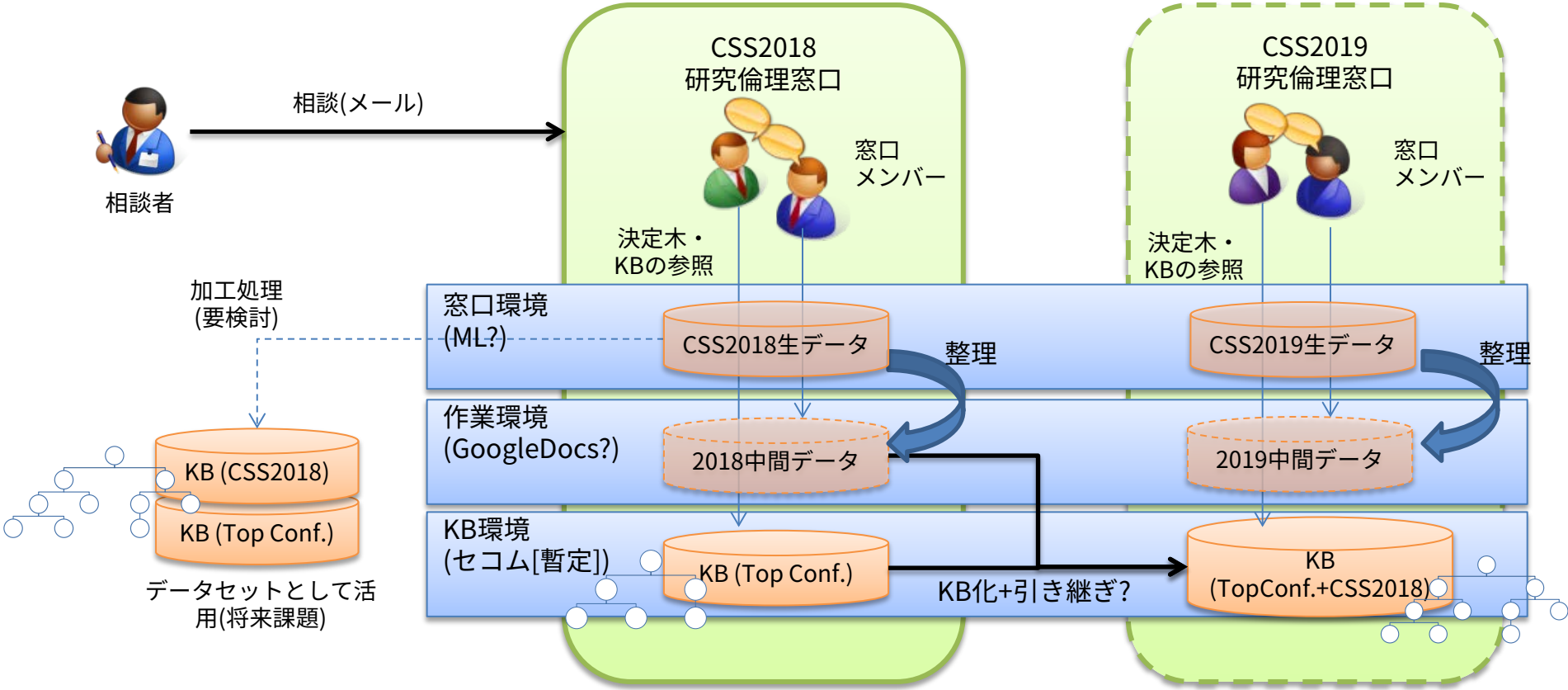
サイバーセキュリティ研究倫理の知識ベース

- 過去5年間の主要国際会議等を対象に研究倫理に関する議論を調査・整理し、知識ベースとして構築した[1]
- 知識ベースは決定木の形で表現可能(詳細後述)
  - 5クラス、20サブクラス、合計207件の条件
  - ユースケースにもとづいて研究の倫理性を判断できる
  - 当該研究に近い過去の事例・議論を容易に発見することができる



[1] 稲垣ら, サイバーセキュリティに関する研究倫理の調査と倫理指針の構築, 2F3-4, 暗号と情報セキュリティシンポジウム2018.

知識ベースの活用案



KBはJSONファイル(決定木で可視化)、生ログは例えばGoogleシート & メールログなどの形式で扱うことを想定(具体的な管理方法については別途議論)

窓口メンバの皆様へのお願い

- 是非知識ベースをご試用いただきご意見ください
  - 全員ではなく興味を持たれた方だけでOKです
- 生データ⇒中間データへの整理についてお認めいただきたく
  - まずは相談情報の蓄積・整理の試行・検討として
  - 整理作業は島岡が負担します
  - 中間データの権利は研究倫理窓口に帰属します

FAQ

- 窓口環境とは具体的に何か？
  - 窓口MLおよび窓口メンバMLとそのメールアーカイブを想定
  - 今年度は明治大学でML運用、メールアーカイブは未確認(ミニマムには各メンバのメール受信環境)
- KB環境はどこで運用するのか？
  - 今年度は弊社で用意します
  - 今年度参照するKBは既知情報のみであり、相談情報は含まれません
  - 次年度以降については後述
- KB環境に必要なスペックは？
  - 知識ベース自体は数百件のルールが記載されたテキストファイル(JSON形式)
  - 現状はPythonの可視化ライブラリとJavaScriptのみで動作しますが可視化のみ
  - 将来的には表示・入力可能とする予定ですが、必要スペックなどはこれから検討していきます
- 作業環境はどこで運用するのか？
  - 相談情報が含まれるため弊社以外の安全な環境が望ましい
  - 現時点ではGoogle Docs(フォーム・スプレッドシート)を検討中
  - Google Docsであればアカウントベースでアクセス制御可能なので来年度以降への引き継ぎもやりやすいと思慮
- CSSのWebサイトで運用できないのか？
  - CSSのWebサイトは、現状では動的コンテンツや実行モジュールを想定していないという認識
  - CSS実行委やCSEC/SPT(/ISEC)にも負担がかかると思慮

サイバーセキュリティ研究倫理のための知識ベースの解説

知識ベース(KB)の概要

- 過去5年分の主要国際会議等962本を対象に研究倫理の議論を調査
  - USENIX Sec, ACM CCS, IEEE S&P, NDSS&UseSec, SOUPS, CREDS, PETS
- Menlo Report, ACM Code of Ethicsなど既存指針も参照
- 倫理に関する議論が含まれる約200本の論文からプラクティスを抽出、KB化
  - SCIS発表時点で未着手だった論文についても鋭意作業中⇒5月中旬までに完了予定
- 現時点で5クラス、20サブクラス、207の条件(と推奨のセット)から成る
  - 若干増える見込み(特に条件)はあるものの、概ね収束
- KBは研究者が自己評価に使うことを想定し、決定木として可視化
  - 現状は決定木は可視化のみで入力機能は未実装
- KBは、決定木として表現しやすいJSON形式で管理されている

KBのクラス・サブクラスと条件数

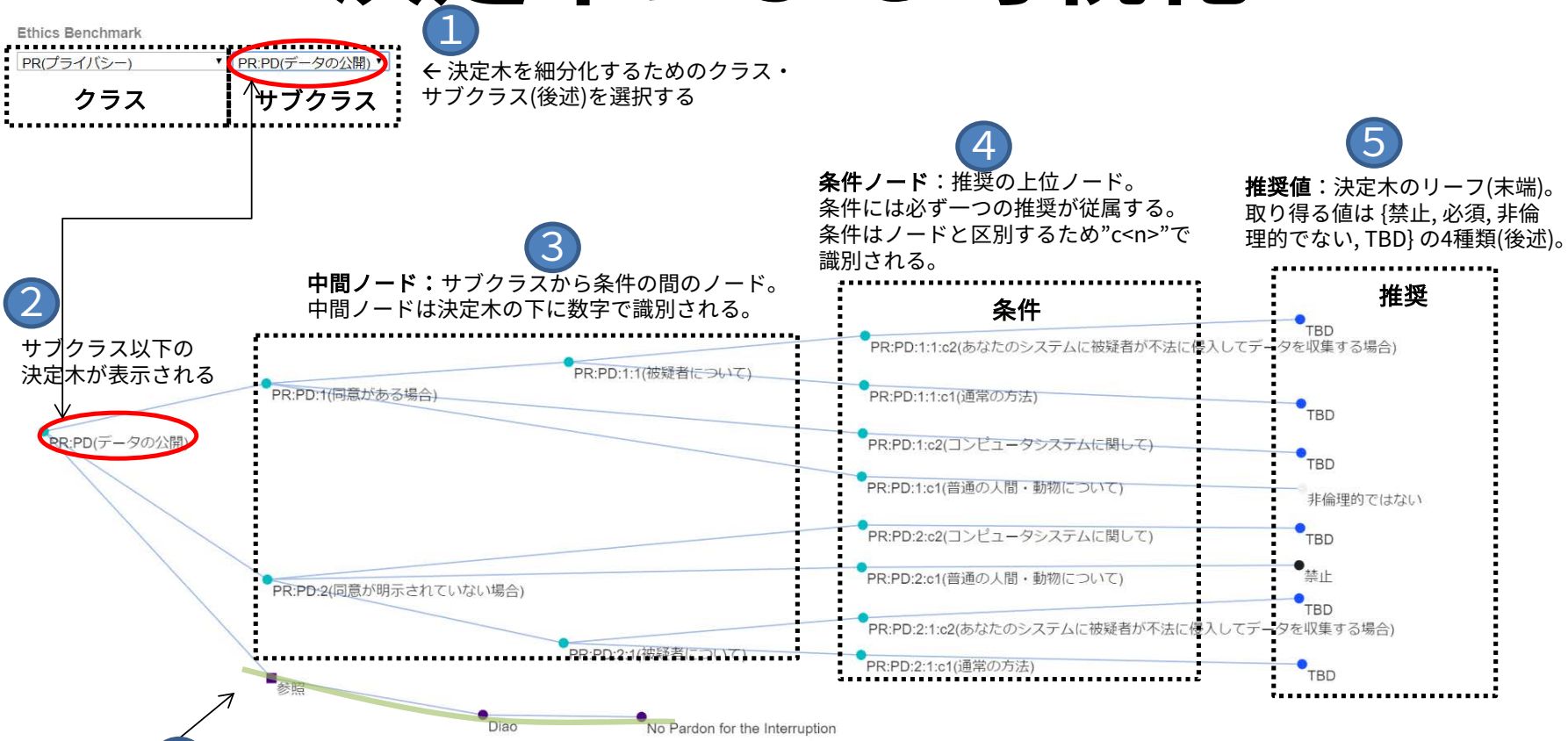
SE:ソフトウェア解析	SE:VR:脆弱性研究	6
	SE:RE:リバースエンジニアリング	4
	SE:MW:マルウェア	11
	SE:DI:開示	23
PR:プライバシー	PR:CD:データ収集	20
	PR:HD:データ処理	24
	PR:PD:データの公開	8
	PR:TD:第三者移転	4
AU:自動化	AU:WS:ウェブスクレイピング・インターネットスキャン	5
	AU:AC:他者のシステムにアクセスすること	11
HS:被験者実験、REBs・IRBs	HS:DE:実験対象の人間または動物を欺く	3
	HS:ML:虚偽などの広告	4
	HS:HP:ハニーポット	6
	HS:CS:犯罪・非倫理的なサービス	2
	HS:EB:REB・IRBに相談すること	4
GR:一般規則	GR:TS:サービス利用規約および関連規約	3
	GR:EC:倫理的な一貫性	10
	GR:DA:文書化と説明責任	3
	GR:CE:ACM Code of Ethics	32
	GR:MR:メンローレポート	25
クラス(5)	**サブクラス(20)**	**条件(207)**

KBのJSONファイル(例)

```
{
  "Category": {
    "SE:ソフトウェア解析": {
            :
      "SE:DI:開示": {
        "参照": {
          "Bhargavan": "On the Practical (In-)Security of 64-Bit Block Ciphers",
          "Bianchi": "What the App Is That?",
              :
        },
        "私的開示": {
              :
          "セキュア通信の利用":{
            "脆弱性がゼロデイの場合": "TBD",
            "脆弱性がエクスプロイトされている場合": "TBD"
          },
        "公開前":{
          "脆弱性がゼロデイの場合": "禁止",
          "脆弱性がエクスプロイトされている場合": "非倫理的ではない"
        },
      :
```

CSSでは、生ログへのポインタを格納可能。
(MLのindexや、もしあれば相談IDなど)
後日KB公開する場合は、例えば投稿論文のDOIなどに
置き換える (既存文献についてもDOI入力中)。

決定木による可視化


Ethics Benchmark
PR(プライバシー)
PR:PD(データの公開)
クラス
サブクラス
1
← 決定木を細分化するためのクラス・
サブクラス(後述)を選択する
2
サブクラス以下の
決定木が表示される
3
中間ノード：サブクラスから条件の間のノード。
中間ノードは決定木の下に数字で識別される。
4
条件ノード：推奨の上位ノード。
条件には必ず一つの推奨が従属する。
条件はノードと区別するため”c<n>”で
識別される。
5
推奨値：決定木のリーフ(末端)。
取り得る値は {禁止, 必須, 非倫
理的でない, TBD} の4種類(後述)。
PR:PD(データの公開)
PR:PD:1(同意がある場合)
PR:PD:1:1(被疑者について)
PR:PD:2(同意が明示されていない場合)
条件
PR:PD:1:1:c2(あなたのシステムに被疑者が不法に侵入してデータを収集する場合)
PR:PD:1:1:c1(通常の方法)
PR:PD:1:c2(コンピュータシステムに関して)
PR:PD:1:c1(普通の人間・動物について)
PR:PD:2:c2(コンピュータシステムに関して)
PR:PD:2:c1(普通の人間・動物について)
PR:PD:2:1:c2(あなたのシステムに被疑者が不法に侵入してデータを収集する場合)
PR:PD:2:1:c1(通常の方法)
推奨
TBD
TBD
TBD
非倫理的ではない
TBD
禁止
TBD
TBD
参照
Diao
No Pardon for the Interruption
6
参照：上位ノード(サブクラス/中間ノード/条件のいずれか)に関連する文献情報。
最も関連するサブクラス・ノード・条件に従属し、下位ノードに文献の著者、タイトルを持つ。

推奨値の説明

1. 禁止(prohibited)：非倫理的な行為
2. 必須(demands)：倫理的であるために不可欠の行為
   - あなたの研究が上位の条件に該当するなら、必須となる。
3. 非倫理的でない(permitted)：非倫理的とは明言できない行為
   - 状況によってはやってもよい可能性がある。判断はケースバイケースになるが、参照ノードがあれば判断材料のひとつになるだろう。
4. TBD(ToBeDetermined): 要決定
   - 先行事例・知見などに乏しく、推奨値を決定できなかったもの。
   - もし推奨値の決定に有益な先行事例・知見などがあれば是非フィードバックいただきたい

作業環境への入力例(1)

倫理指針の決定木を用いた自己評価結果

決定木を使って、あなたの研究行為に該当する条件を探してください。
該当する条件の推奨値が、あなたの研究行為と一致するかどうかについて、以下に回答してください。

○：あなたの研究行為について、推奨値(と一致する条件番号を回答してください。あなたの研究行為に該当しない条件については回答不要です。

解説：推奨値とあなたの研究行為の関係が以下の場合に○を回答してください。
- 必須: あなたの研究行為が条件を実施しているなら○
- 禁止: あなたの研究行為が条件を実施していないなら○
- 非倫理的ではない: あなたの研究行為が条件を実施しているなら○
- TBD: あなたの研究行為が条件を実施しているなら○

カンマで○の条件番号を入力してください（例:“PR:HD:1:2:c1, SE:DI:2:c3”)。

カンマで○の条件番号を入力してください（例:“PR:HD:1:2:c1, SE:DI:2:c3”)。

回答を入力

作業環境への入力例(2)

×：あなたの研究行為について、推奨と一致しない条件番号を回答してください。あなたの研究行為に該当しない条件については回答不要です。

解説：推奨値とあなたの研究行為の関係が以下の場合に×を回答してください。
- 必須: あなたの研究行為が条件を実施していないなら×
- 禁止:あなたの研究行為が条件を実施している/してしまったなら×

TBD, 「非倫理的ではない」は回答できません。

カンマで×の条件番号を入力してください（例:“PR:HD:1:2:c1, SE:DI:2:c3”)。

回答を入力

回答のコピーを自分宛に送信する

送信

Google フォームでパスワードを送信しないでください。

作業環境から窓口メンバへの出力案

	著者１のID	著者２のID
メールアドレス	dareka@kaisya.co.jp	
タイムスタンプ	2018-06-14	
○	PR:PD:1:1:C1 . .	
✖	PR:PD:2:C1 . .	
参照	URL/DOI/ポインター	

著者１の倫理評価

○
PR:PD:1:1:C1
[プライバシー]データ公開について、被疑者の同意がある場合に、通常の方法で同意を取得している。
Ref: [PR:PD], <著者>, “タイトル”, <DOI>

✖
PR:PD:2:C1
[プライバシー]
禁止：(普通の人間・動物について、同意が明示されていないデータを公開することは禁止されている。)
Ref: [PR:PD], <Diao>, “No Pardon for the Interruption”, <DOI>

当該条件から一番近い位置の参照

Googleフォーム等の集計結果から、human-readableな倫理評価の結果をスクリプトで生成する