

医療合成データ作成・攻防コンテスト

相手顧客に競合の欠陥を暴け!

安全で有用な合成データを作れ!

ライバル企業の弱点を探せ!

エントリー期間
2026. 7/30(木) ~ 9/7(月)

「エントリー可能な方」CSS2026 参加者さま
エントリーはQRコードから

コンテスト 予備選
2026. 8/9(日) ~ 9/7(月)

コンテスト 本選
2026. 9/12(土) ~ 10/11(日)

コンテスト 発表・表彰式
下記日曜のいずれか
2026. 10/19(日) ~ 10/23(金)
アクティシティ浜松

主催 情報処理学会コンピュータセキュリティ研究会、PWS組織委員会

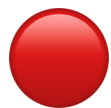
PWS Cup 2026 説明会

2026年7月30日(木) 18:15~19:00

オンライン (zoom)

PWS Cup 2026 WG長 小野元

注意



- 本説明会は録画しており、説明会終了後録画を公開します
- チャットや発言も公開します
- Q&Aタイムで質問する方は留意してください
 - 説明会に参加していたことが第三者にもわかる
 - 踏み込んだ質問をすると作戦が他チームにバレる
- このスライドも録画とともに公開します

説明会の流れ

- 諸注意 ← 今ここ
- PWS Cup 2026の説明（約30分）
- Q&A（約15分）19時には終了
 - メール等で随時受付
 - 詳細なルールブックは別途公開

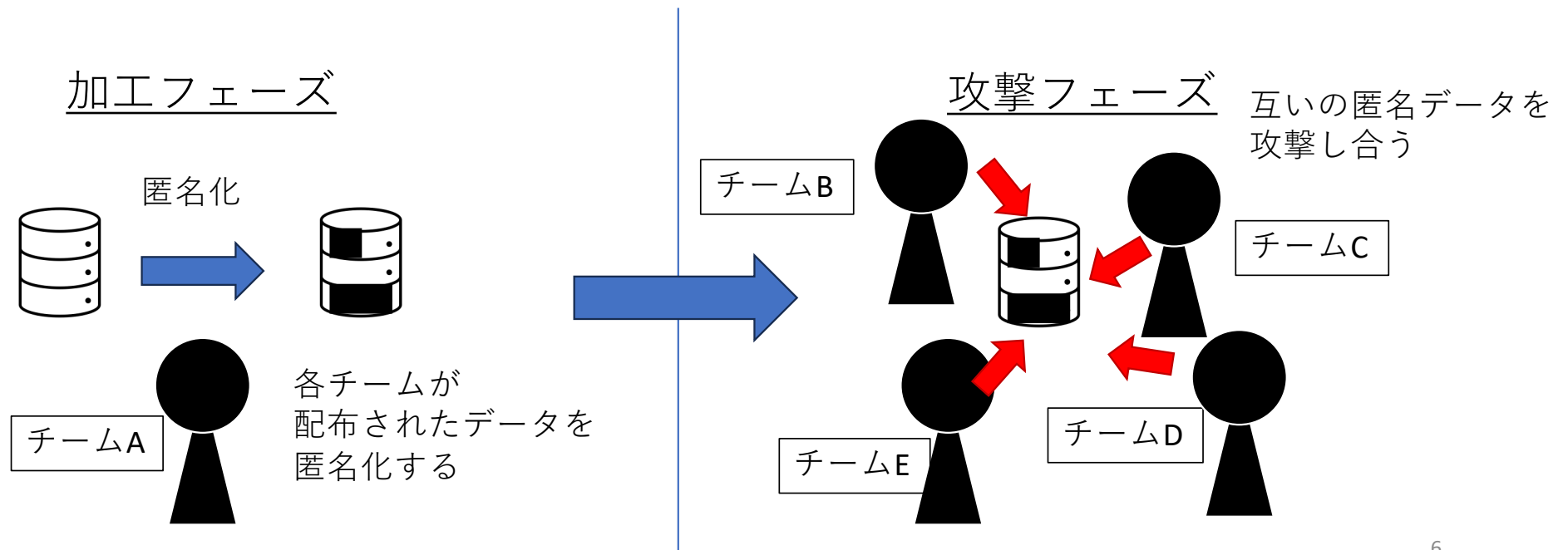
概要と全体の流れ

PWS Cupとは

- 2015年から続くデータ匿名化コンテスト
- **防御と攻撃**両方を実施する**チーム戦**
- 有用性得点と匿名性得点を競う
- 加工フェーズ：有用性を保ちながらデータを匿名化
- 攻撃フェーズ：別チームの匿名データを攻撃
 - 匿名性得点: 別チームの攻撃が成功しづらいほど高得点
 - 攻撃を頑張る = 別チームの匿名性得点を下げる
- **総合力**が問われる競技性 → 力試し、教育の場として継承

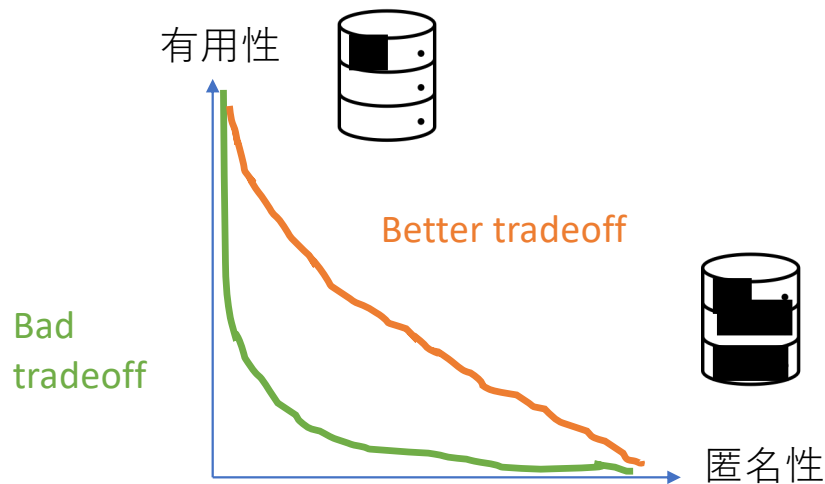
PWS Cupの流れ

- 加工フェーズ→攻撃フェーズ
- 予備戦と本戦で合計4フェーズで戦う



匿名性得点と有用性得点を競う

- 匿名性(防御力)と有用性をバランスする加工方法・程度を探索
- 敵チームの得点を下げるために効果的な攻撃を探索



匿名性を向上させると
有用性は下がる
(元データからの乖離が大きくなる)

どのくらい下がるかは
加工方法によって異なる

PWS 2026のコンテストストーリー

- あなたのチームは、健診データを匿名化して売るスタートアップ企業。手元にあるのは健診記録 B——年齢・血圧・血糖値から、10年の追跡で糖尿病を発症したか・いつ発症したかまで並んだ個人データです。個人が特定されない匿名化データ C を作って売ります。
- 顧客は研究者。ひとは**集団全体の傾向**を知りたい**疫学研究者**、もうひとは数少ない**非典型患者**（MODY・1型）を調べたい**研究者**。集団の傾向も、発症の要因も、希少な患者の姿も——どれかを欠けば顧客の誰かが不満を持ちます。これが有用性（U）。
- 同業他社（他チーム）はライバル。あなたの C を粗探しし、実在の個人を当てられれば（メンバーシップ推論・属性推論）信用を失います。攻撃されても個人が割れないこと＝匿名性（Anon）。あなたも攻撃側に回ります。

事務局の狙い：参加＝研究

- 参加経験がそのまま研究につながるように設計
 - 新しい論文を参照して問題設定・評価指標を工夫
 - 見込みのある防御手法・攻撃手法は積極的に情報共有
- 卒論や学会発表につながる事が理想
 - 学会などの情報を適宜共有

スケジュール

- 参加受付開始：説明会后～
- 練習（CodaBench）：7月30日(木)－8月8日(土)
- 予備戦
 - 加工フェーズ：8月09日(日)－8月22日(土)
 - 攻撃フェーズ：8月25日(火)－9月07日(月)
- 本戦
 - 加工フェーズ：9月12日(土)－9月25日(金)
 - 攻撃フェーズ：9月28日(月)－10月11(日)
- 発表会・表彰式：CSS2026会期中(10月19日(月)～23日(金)いずれか)
 - 対面

約2ヶ月の長丁場

データと採点

詳細はルールブック参照

特に重要と思われる部分と理念をお話しします

データ

- 1チームに配るコホート B は 1,049 人 × 14 列
 - 列の内訳：属性・健診値・アウトカム (time・onset・death)
 - 行の内訳：自チームだけの行 313人／他チームの一部と共有 736人
 - 転職で複数の健康保険会社にデータがある人のイメージ
- 希少な患者（非典型糖尿病）は 24 人 = 2.3%
 - MODY（HNF1A 型）17 人・1型糖尿病 7 人
 - 人数も型の内訳も、全チームで厳密に同じ
- 追跡は最長10年
 - 10年のうちに発症する人はおよそ1割・死亡は7～8%
 - 希少群だけは 3～4割が発症
- 用意したコホートは30セット = 参加できるチーム数の上限

データのイメージ（練習用から抜粋）

年齢	性別	県	BMI	SBP	FPG	喫煙	追跡年	発症	死亡
44	男	神奈川県	25.8	116	113	0	10.0	0	0
40	男	大阪府	30.8	145	120	0	1.3	1	0
40	男	神奈川県	21.4	93	114	0	2.9	1	0
73	男	愛知県	24.4	145	119	1	2.5	0	1
45	男	福岡県	35.5	161	125	1	0.3	1	0
58	女	奈良県	21.9	143	100	0	10.0	0	0

実際に配られる CSV の一部（record_id・TG・HDL・ALT などを含め全14列）。SBP=収縮期血圧, FPG=空腹時血糖。喫煙・発症・死亡は 1=あり／0=なし。

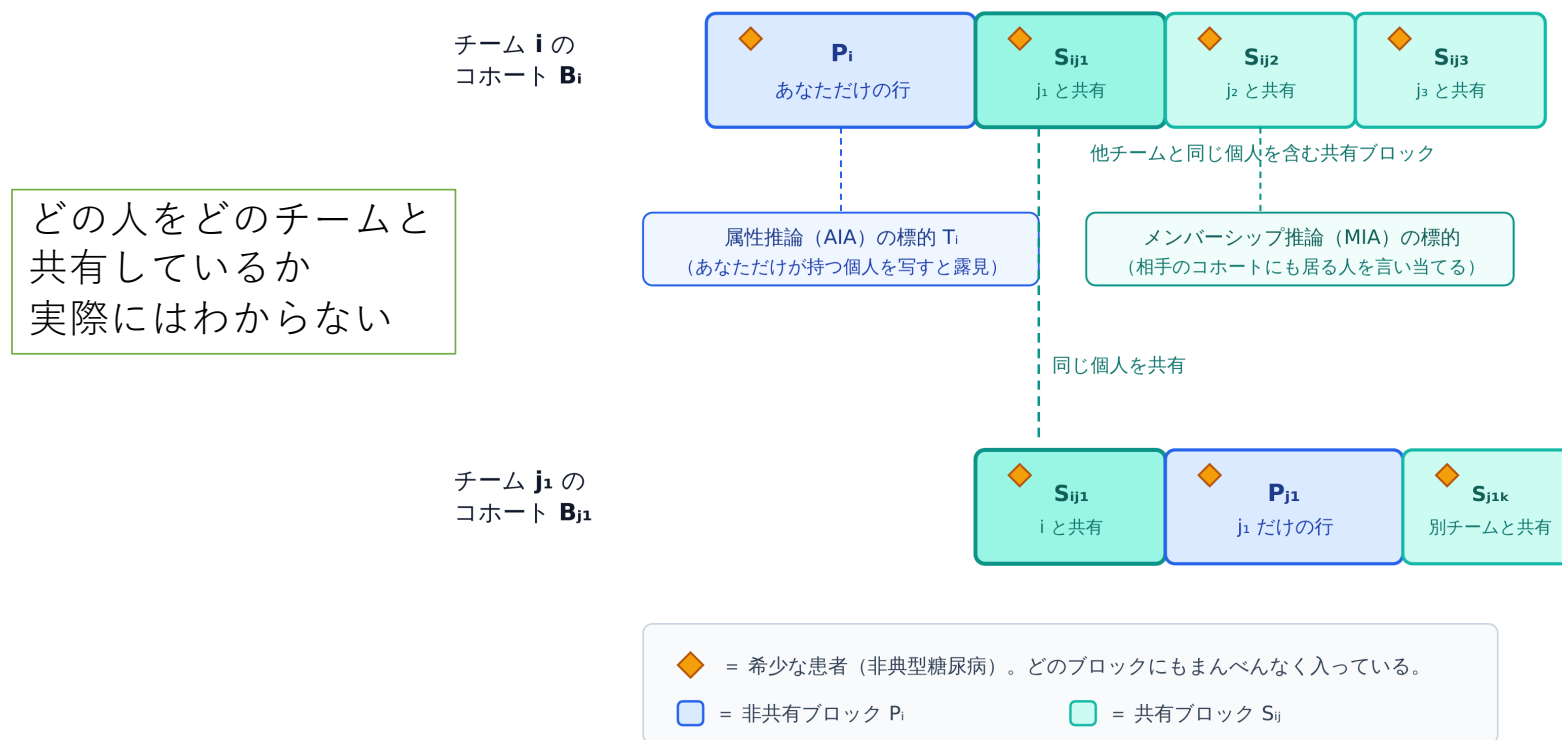
▶ 「発症したか・いつ発症したか」 — “特定タスクの忠実さ”で効いてくる列。

▶ 色の付いた行は希少タイプ（若くして急に発症・やせ型など。全体の2.3%） — ここを潰さず残せるかが問われる。

チームごとのデータの偏りと共有

図2：コホート B_i の構造

1つの母集団 A から切り出した、あなたのチームのコホート B_i 。性質の違う2種類の行でできている。



得点 $\min(U, \text{Anon})$

- 有用性 U : もとデータの性質をどのくらい引き継ぐか評価。
引き継いでいるほど顧客が喜ぶ
- 匿名性 Anon : 他チームからの攻撃への堅牢性を評価。
堅牢なほど、告げ口が無い
- どちらを捨てても高得点にはならない

有用性 $U = \min(U_{\text{gen}}, U_{\text{spec}}, U_{\text{rare}}, U_{\text{valid}})$

記号	名前	何を測るか
U_{gen}	一般有用性 (gen = general)	データ全体の分布が本物にどれだけ似ているか
U_{spec}	特定タスクでの結果忠実度 (spec = specific)	「何が発症に効くか」の分析結果が本物と同じ結論になるか
U_{rare}	希少群の有用性 (rare)	希少群（非典型糖尿病）を潰さず再現できているか
U_{valid}	妥当性 (valid = validity)	数値だけ取り繕う小細工を防ぐ最低限の整合チェック

一部の評価を捨てる作戦に高得点を与えない設計

匿名性 $\text{Anon} = 1 - \max(a_{\text{MIA}}, a_{\text{AIA}})$

- MIA成分 a_{MIA} : Membership inference attackが
どのくらい成功したかを評価する量
 - 各敵チームが自チームと共有しているのはどの人が当てる
- AIA成分 a_{AIA} : 属性推論攻撃(attribute inference attack)が
どのくらい成功したかを評価する量
 - 標的の一部の情報（年齢・血圧・県など）だけから隠れた発症時期を推定する攻撃。
 - 母集団しか見なくても当たる分を超えたら失点

表彰

- 競技で決まる賞 = 結果を評価
 - 総合順位 1～5位
 - ベストアタック賞
 - 敵チームの匿名性を一番多く下げたチームを表彰
- 発表会で決まる賞 = 過程を評価
 - オープンサイエンス賞
 - コミュニティに有益な情報を共有してくれたチームを表彰
 - 特別審査員賞
 - 医療データという文脈で良いプレゼンをしたチームを表彰

参加方法

参加準備

- メンバーを集める
 - 1人でも参加OK
- ルールブックを読む
- 代表者がCodaBenchのアカウントを作成
 - もう持っていれば新規作成は不要
 - <https://www.codabench.org/>
 - <https://www.codabench.org/competitions/17698/>
 - コンペティションに参加(My submission -> register)
- 代表者がwebフォームに入力して参加登録
 - PWS Cup 2026 webサイトから
 - 参加登録後、事務局がCodaBenchで参加を承認 ← ここで完了

CSS2026への参加

- 発表会・授賞式は**CSS2026**の**1セッション**として開催
- 当日少なくとも一人は発表会に参加してください
 - 現地参加が難しいチームは要相談
- 5～10分程度のプレゼン
- ポスターセッションへの参加
- 賞状・記念品の贈呈

匿名参加OK

- PWS Cupの伝統
 - 事務局には参加者の氏名と所属を伝えてもらう
 - 一般公開しないことは可
-
- 企業人に参加してもらいやすくする配慮
 - 例年、受賞が確定してから公開する参加者も

メンバーは原則5人以下

- チーム間の人数差が多すぎると不公平感が出る
- 自分の意見を作戦に組み込んでもらえない人が出てくる
- 例外：学生主体のチームは5+2まで
 - 責任者(必須)：**助言**と**監督**を行う。准教授や教授を想定
 - 指導者(任意)：技術的な**助言**を行う。助教や外部メンターを想定
 - 複数チームの指導を兼務しても良いが、節度を持って助言すること
 - 兼務する場合は**氏名と所属を公開** → 公平性を強く意識すること
 - 公平性の調整と学生の教育効果の最大化のための措置

チーム編成に関する注意

- 参加登録後のチームメンバー変更不可
- 4人以下のチームへの完全新規のメンバー追加はワーキンググループで審議
- 今回のルールは「複数チームのデータBを知っている」ことがアドバンテージになってしまうため
- 「5人チームだったけど、4人蒸発した」は救済できない
- 誰とチームを組むか登録前によく考えること

チーム登録時に申告をお願いします

- メンターやチーム合併の斡旋
 - 事務局で努力します。確約はできません
- 発表会・授賞式でのアクセシビリティのご相談
 - 手話通訳など学会側の負担で手配可能なサービスもあります

代表者はメールを確認してください

- 競技用データの配布はメールで情報伝達します
- トラブル等があった場合は事務局からメールをします
 - サーバーダウン
 - スケジュール変更
 - 採点プログラムのバグ修正

その他

- LLMの使用は特に制限しない
 - 発表会でどんなLLMをどの程度使ったか報告してほしい
 - ハッキングなどの不正はしないこと
- 不正をしない・させない
 - 他チームのもとデータを見ない
 - 他チームから見える場所にデータを保管しない
 - 指導者・責任者も注意！

リンク集

- CodaBenchの競技ページ
 - <https://www.codabench.org/competitions/17698/>
- スターターキット
 - ルールブック・練習データ・参照実装・教材ノートブック
 - <https://github.com/pwscup/pwscup2026-kit>
- ルールブック
 - https://github.com/pwscup/pwscup2026-kit/blob/main/rulebook/PWS_Cup2026_%E3%83%AB%E3%83%BC%E3%83%AB%E3%83%96%E3%83%83%E3%82%AF_orientation-20260730c.pdf

Q & A

チャット、音声受け付けます

- メールでも受け付けます
- pwscup2026-info@csec.ipsj.or.jp
 - [加筆] メール以外の経路でも対応いたします
- CodaBenchの不具合、採点プログラムのバグが有り得ます
 - あれ？と思ったら事務局にご相談ください
 - 競技前・中なら対応可能です
- 録画、スライド、ルールブックはWebで