PWS Cup セッション

PWS 実行委員会

本セッションについて

- PWS Cup:
  匿名化(プライバシー保護)と
  それを暴く攻撃の
  技術を競うチーム対抗戦

- PWS Cup 2023 のスケジュール:

8月17日	募集開始	
8月29日～9月21日	予備戦(1周目)	} オンラインで
9月28日～10月24日	本戦(2周目)	} 2周実施済み
11月1日	結果発表, 参加チームによる手法の紹介	

本日の予定

- 09:25 – 09:55 開会
  - 開会，説明，結果発表
- 09:55 – 11:10 プレゼン発表

---------- 休憩(20分) ----------

- 11:30 – 12:20 ポスター発表(前半)

---------- 昼休み (80分) ----------

- 13:40 – 14:30 ポスター発表(後半)

---------- 休憩・審査 (10分) ----------

- 14:40 – 15:10 閉会
  - 発表賞の発表，審査員講評，記念撮影，閉会

参加チーム紹介

チームID	チーム名	チーム代表者	所属
01	堅あげポテト	-	-
02	神馬大 νττ	小林雅弥	神奈川大学
03	ぼっちくんち	井口誠	Kii株式会社
04	宮地研.exe	山月達太	大阪大学
05	F.SE	-	-
06	GOUTAI	伊藤聡志	-
07	Guardians of Privacy vol.23	小嶋柊偉	静岡大学
08	ミズホノデシ	-	静岡大学
09	みずほに恩返し	-	静岡大学
10	ステテコ清水	清水正浩	明治大学
11	NYCU Quester	-	-
12	タケタケ	-	-
13	TCFSH	Chen Bo Hsuan	-
14	Enthusiasm	-	-
15	鋼鉄の錬金術師	中田良祐	日鉄ソリューションズ

PWS Cup 2023 競技ルール

PWS Cup 2023 で想定する世界

登場人物

- **加工者**:
  自身の持つデータを
  加工して公開

元のデータ

名前	年代	地域	服
Alice	30代	近畿	S
Bob	40代	九州	M
Chris	40代	関東	M
Dave	20代	九州	L

⇨

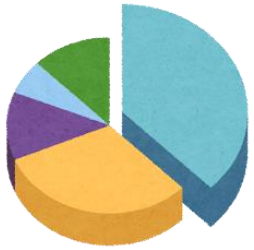
加工データ

年代	地域	服
30代	近畿	S
40代	九州	M
40代	関東	M
20代	九州	L

- **活用者**:
  加工データを分析

年代	地域	服
30代	近畿	S
40代	九州	M
40代	関東	M
20代	九州	L

⇨



- **攻撃者**:
  加工データを使って
  知人の秘密を暴く
  (**属性推定攻撃**)

知人の断片情報

名前	年代	地域	服
Alice	30代	近畿	?

＋

年代	地域	服
30代	近畿	S
40代	九州	M
40代	関東	M
20代	九州	L

➡ **Aliceの服はS！**

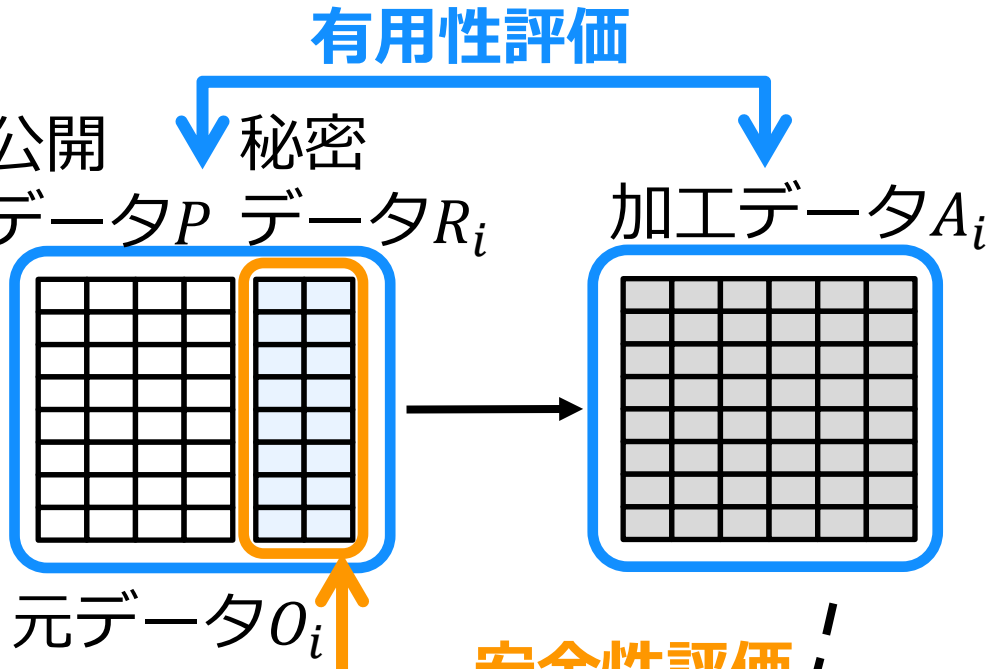
コンテスト概要

- 加工対象のデータは1人1レコードの**表形式**
- 参加チームは**加工者**または**攻撃者**となる(兼任可能)
- コンテストは**2段階**で実施

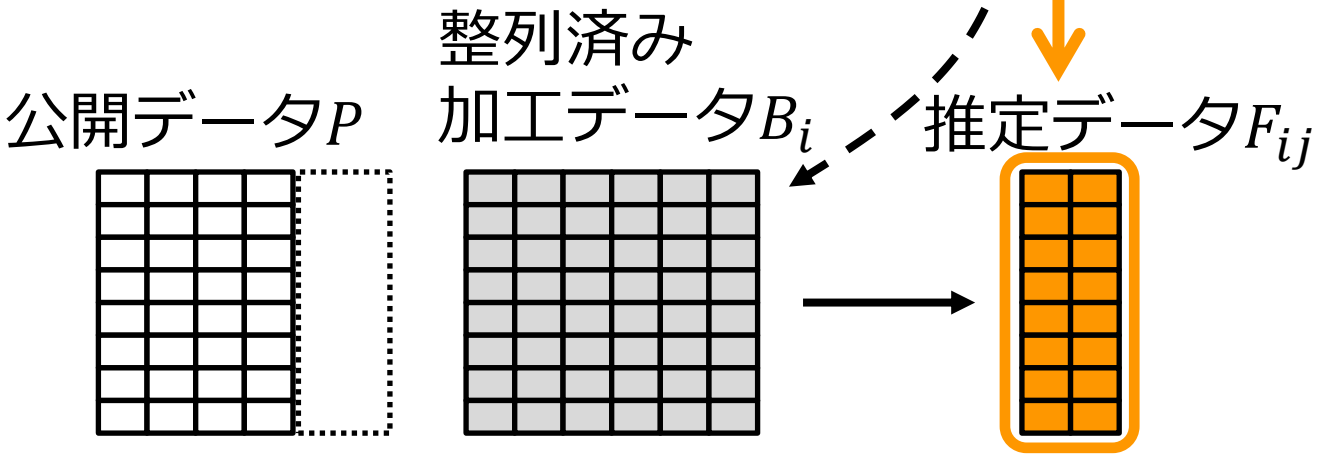
1. **加工段階**
   加工者は与えられたデータを**加工**して提出
2. **攻撃段階**
   攻撃者は各加工データに**属性推定**で攻撃

コンテストの流れ

1. 加工段階
(チームiが加工)



2. 攻撃段階
(別チームjが
チームiを攻撃)



データセット

- 1行1顧客の表
- 各セルは0から9の10種類のカテゴリ値
- セルごとに独立に，ランダムに生成
  - 公開データは小さい数字ほど出やすい
  - 秘密データは一様ランダム

公開データP

年代，居住エリア，…

0,8,9,5,0,0
1,2,6,1,2,4
3,0,4,2,4,7
7,0,1,7,0,6
3,2,0,6,1,1
0,1,1,8,3,0
0,7,3,2,5,1

秘密データR_i

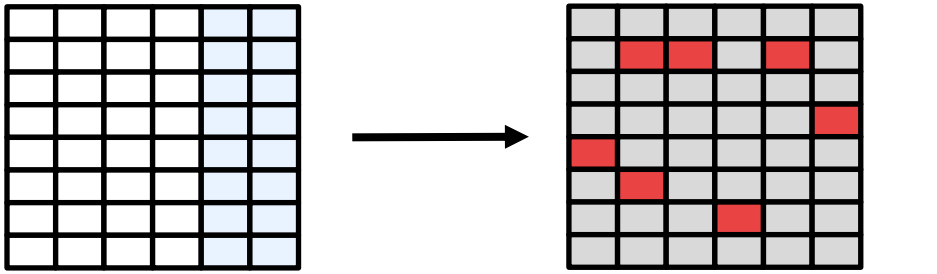
(離散化した)体重，年収，…

8,8,6,2,8,7,2,1,5,4
4,5,7,3,6,4,3,7,6,1
3,5,8,4,6,3,9,2,0,4
2,4,1,7,8,2,9,8,7,1
6,8,5,9,9,9,3,0,0,2
8,8,2,9,6,5,6,6,6,3
8,2,1,4,8,1,6,9,5,1

有用性評価

有用性評価値 = (変更しなかったセルの割合)

元データO_i　　加工データA_i

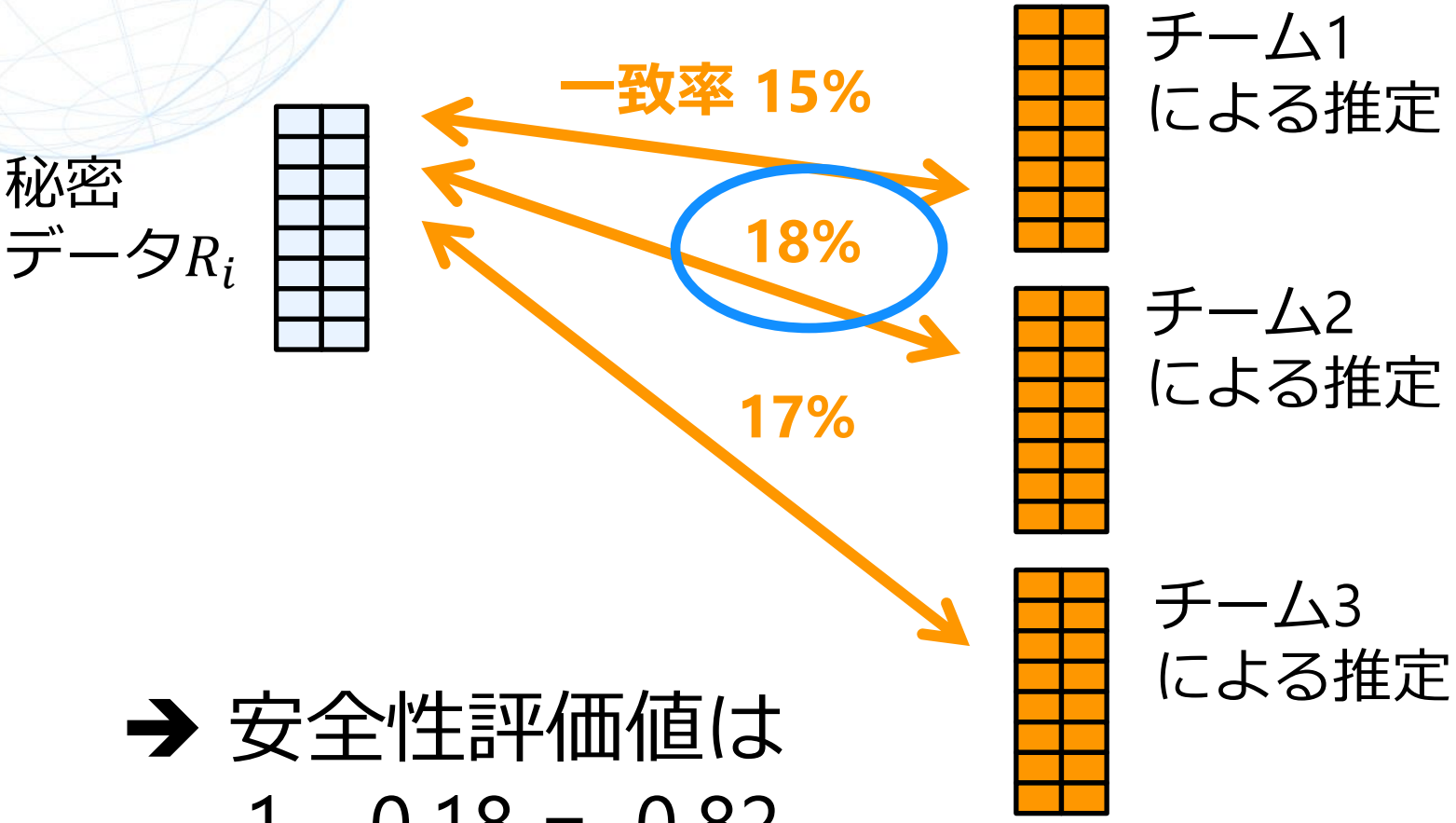


15%のセルを変更

➔ 有用性評価値は 1 - 0.15 = 0.85

安全性評価

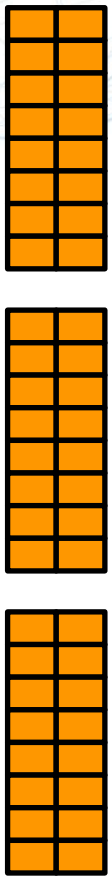
安全性評価値 := 1 – (最高の一致率)



➔ 安全性評価値は
1 – 0.18 = 0.82

攻撃の評価

攻撃評価値 = 全チームに対する推定の一致率の平均値



チーム1
に対する推定 → 一致率10%

チーム2
に対する推定 → 一致率20%

チーム3
に対する推定 → 一致率18%

攻撃評価値は
(0.1+0.2+0.18)/3
= 0.16

各チームの評価

- 加工部門の順位:
  有用性評価値 + 安全性評価値 が大きい順
  (加工セル割合) (1-最悪の一致率)

- 攻撃部門の順位:
  攻撃評価値 が大きい順
  (全チームへの攻撃の一致率平均)

最終的な順位の決定方法

実際のコンテストは
重み 1 : 9 で予備戦と本戦の2周実施

➔ 最終的な順位は以下のように決定

- **全体加工順位**:
  (予備戦加工順位) + (本戦加工順位) × 9 の昇順
- **全体攻撃順位**:
  (予備戦攻撃順位) + (本戦攻撃順位) × 9 の昇順
- **総合順位**:
  (全体加工順位) + (全体攻撃順位) の昇順

パラメータ設定

公開データP

秘密データR_i

行数 10^5

5,8,9,5,0,0	8,8,6,2,8,7,2,1,5,4
1,7,6,9,2,4	4,5,7,3,6,4,3,7,6,1
5,2,4,2,4,7	3,5,8,4,6,3,9,2,0,4
7,9,1,7,0,6	2,4,1,7,8,2,9,8,7,1
9,9,7,6,9,1	6,8,5,9,9,9,3,0,0,2
0,1,8,8,3,9	8,8,2,9,6,5,6,6,6,3

列数 6 (予備戦) or 8 (本戦)

列数10

コンテストの結果

予備戦の結果(攻撃一致率他)

NTT

攻撃チーム

加工チーム	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	安全性	有用性
01	0.000	0.155	0.021	0.021	0.021	0.021	0.000	0.000	0.000	0.021	0.021	0.021	0.101	0.021	0.000	0.845	0.921
02	0.145	0.000	0.153	0.154	0.102	0.155	0.153	0.000	0.000	0.153	0.100	0.153	0.100	0.100	0.000	0.845	0.899
03	0.148	0.207	0.000	0.215	0.184	0.159	0.198	0.000	0.000	0.175	0.100	0.193	0.100	0.100	0.000	0.785	0.905
04	0.103	0.124	0.103	0.000	0.100	0.103	0.100	0.000	0.000	0.102	0.100	0.102	0.100	0.100	0.000	0.876	0.846
05	0.101	0.100	0.100	0.100	0.000	0.100	0.100	0.000	0.000	0.100	0.100	0.101	0.100	0.100	0.000	0.899	0.874
06	0.100	0.101	0.101	0.100	0.101	0.000	0.100	0.000	0.000	0.100	0.101	0.100	0.100	0.101	0.000	0.899	0.697
07	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000
08	0.649	0.769	0.761	0.769	0.653	0.760	0.162	0.000	0.000	0.761	0.103	0.761	0.100	0.103	0.000	0.231	0.990
09	0.376	0.513	0.407	0.127	0.130	0.494	0.099	0.000	0.000	0.495	0.101	0.100	0.100	0.101	0.000	0.487	0.909
10	0.815	0.821	0.815	0.821	0.814	0.815	0.000	0.000	0.000	0.000	0.816	0.815	0.100	0.816	0.000	0.179	0.997
11	0.128	0.165	0.105	0.165	0.122	0.120	0.100	0.000	0.000	0.116	0.000	0.115	0.100	0.100	0.000	0.835	0.875
12	0.379	0.516	0.495	0.127	0.486	0.496	0.100	0.000	0.100	0.498	0.100	0.000	0.100	0.100	0.000	0.484	0.909
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000
14	0.184	0.252	0.231	0.235	0.101	0.232	0.100	0.000	0.000	0.233	0.100	0.100	0.100	0.000	0.000	0.748	0.812
15	0.112	0.137	0.108	0.111	0.110	0.111	0.100	0.000	0.000	0.111	0.100	0.110	0.100	0.100	0.000	0.863	0.875
攻撃(平均)	0.374	0.419	0.386	0.353	0.352	0.397	0.165	0.143	0.150	0.348	0.274	0.334	0.164	0.274	0.143		

- 安全性，有用性ともに0.9近くのチームもある
- 01に対しては0.1 (でたらめに攻撃した場合の期待値)を大きく下回る成功率の攻撃が多い

本戦の結果(攻撃一致率他)

攻撃チーム

加工チーム	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	安全性	有用性
01	0.000	0.067	0.100	0.091	0.072	0.100	0.100	0.100	0.055	0.100	0.100	0.073	0.100	0.100	0.000	0.900	0.825
02	0.109	0.000	0.100	0.176	0.151	0.111	0.100	0.100	0.101	0.100	0.127	0.100	0.099	0.100	0.000	0.824	0.812
03	0.099	0.143	0.000	0.147	0.120	0.118	0.100	0.100	0.110	0.100	0.116	0.117	0.105	0.101	0.000	0.853	0.780
04	0.130	0.142	0.118	0.000	0.100	0.115	0.100	0.100	0.101	0.101	0.120	0.115	0.101	0.100	0.000	0.858	0.845
05	0.114	0.152	0.121	0.153	0.000	0.126	0.104	0.100	0.121	0.100	0.124	0.154	0.125	0.100	0.000	0.846	0.846
06	0.100	0.141	0.107	0.141	0.141	0.000	0.106	0.100	0.100	0.100	0.118	0.108	0.101	0.100	0.000	0.859	0.774
07	0.213	0.309	0.225	0.358	0.280	0.280	0.000	0.100	0.276	0.100	0.321	0.293	0.101	0.100	0.000	0.642	0.913
08	0.100	0.412	0.116	0.101	0.357	0.124	0.100	0.000	0.111	0.100	0.412	0.124	0.268	0.100	0.000	0.588	0.787
09	0.100	0.171	0.156	0.100	0.100	0.156	0.100	0.100	0.000	0.100	0.100	0.100	0.100	0.101	0.000	0.829	0.711
10	0.167	0.234	0.076	0.161	0.173	0.078	0.101	0.100	0.072	0.000	0.101	0.079	0.162	0.100	0.000	0.766	0.877
11	0.172	0.246	0.130	0.245	0.178	0.181	0.100	0.100	0.100	0.100	0.000	0.237	0.101	0.100	0.000	0.754	0.890
12	0.100	0.251	0.106	0.168	0.220	0.106	0.107	0.100	0.106	0.100	0.113	0.000	0.100	0.100	0.000	0.749	0.816
13	0.191	0.058	0.066	0.268	0.267	0.100	0.100	0.099	0.062	0.100	0.213	0.131	0.000	0.100	0.000	0.732	0.893
14	0.106	0.823	0.655	0.871	0.871	0.815	0.100	0.099	0.100	0.100	0.874	0.742	0.871	0.000	0.000	0.126	0.828
15	0.113	0.191	0.131	0.147	0.155	0.156	0.100	0.100	0.131	0.099	0.100	0.153	0.101	0.100	0.000	0.809	0.846
攻撃(平均)	0.130	0.239	0.158	0.223	0.228	0.183	0.101	0.100	0.110	0.100	0.210	0.180	0.174	0.100	0.000		

- 安全性，有用性ともに0.85程度のチームもある
- 01, 10, 13 には 0.1 を大きく下回る成功率の攻撃も

順位

チームID	加工部門順位			攻撃部門順位				
	予備戦	本戦	全体	予備戦	本戦	全体	加+攻	総合順位
01	2	1	1	4	9	9	10	5
02	3	7	6	1	1	1	7	3
03	7	8	8	3	8	8	16	7
04	5	2	2	5	3	3	5	1
05	1	3	3	6	2	2	5	2
06	8	9	9	2	5	5	14	6
07	14	12	12	11	11	11	23	12
08	12	14	14	14	14	14	28	15
09	10	13	13	13	10	10	23	13
10	13	6	7	7	13	13	20	11
11	6	5	5	10	4	4	9	4
12	11	11	11	8	6	6	17	9
13	14	10	10	12	7	7	17	8
14	9	15	15	9	12	12	27	14
15	4	4	4	14	15	15	19	10

入賞チーム (CSS表彰式で表彰)

	総合	加工部門	攻撃部門
優勝	宮地研.exe (チームID: 04)	堅あげポテト (チームID: 01)	神馬大 νττ (チームID: 02)
2位	F.SE (チームID: 05)	宮地研.exe (チームID: 04)	F.SE (チームID: 05)
3位	神馬大 νττ (チームID: 02)	F.SE (チームID: 05)	宮地研.exe (チームID: 04)

優秀発表賞: ？？？(本セッションの発表で決定)

発表賞の説明

優秀発表賞について

- 審査基準

コンテストで使った手法をわかりやすく伝え，
聴衆の匿名化・属性推定技術の向上に
最も寄与したと考えられる発表

- 選定方法
  - ポスター発表終了後に，各審査員一票ずつで，優秀発表賞にふさわしいと考えるチーム一つに無記名投票
  - 最多得票のチームが受賞
    - 複数存在の場合は審査員長がその中から選択

優秀発表賞の審査員

お名前(五十音順・敬称略)	ご所属
須川賢洋	新潟大学
高橋翼	LINEヤフー
玉井睦	セコム
村上隆夫*	統計数理研究所
南和宏	統計数理研究所

* 審査員長

黄色いタグが目印です

プレゼン発表の進め方

- 各チームの発表時間は5分です
(交代の時間を含む)
- 5分経過した場合には合図をするので,
直ちに終了してください

優秀発表賞について

- 審査基準

コンテストで使った手法をわかりやすく伝え，
聴衆の匿名化・属性推定技術の向上に
最も寄与したと考えられる発表

- 選定方法
  - ポスター発表終了後に，各審査員一票ずつで，優秀発表賞にふさわしいと考えるチーム一つに無記名投票
  - 最多得票のチームが受賞
    - 複数存在の場合は審査員長がその中から選択

優秀発表賞

堅あげポテト

(チームID: 01)

審査員講評

お名前(五十音順・敬称略)	ご所属
須川賢洋	新潟大学
高橋翼	LINEヤフー
玉井睦	セコム
村上隆夫*	統計数理研究所
南和宏	統計数理研究所

* 審査員長

入賞チーム (CSS表彰式で表彰)

	総合	加工部門	攻撃部門
優勝	宮地研.exe (チームID: 04)	堅あげポテト (チームID: 01)	神馬大 νττ (チームID: 02)
2位	F.SE (チームID: 05)	宮地研.exe (チームID: 04)	F.SE (チームID: 05)
3位	神馬大 νττ (チームID: 02)	F.SE (チームID: 05)	宮地研.exe (チームID: 04)

優秀発表賞: 堅あげポテト

副賞: 八女提灯 (東さんセレクト)

デザインはお楽しみ♪

入賞チーム (CSS表彰式で表彰)

	総合	加工部門	攻撃部門
優勝	宮地研.exe (チームID: 04)	堅あげポテト (チームID: 01)	神馬大 νττ (チームID: 02)
2位	F.SE (チームID: 05)	宮地研.exe (チームID: 04)	F.SE (チームID: 05)
3位	神馬大 νττ (チームID: 02)	F.SE (チームID: 05)	宮地研.exe (チームID: 04)

優秀発表賞: 堅あげポテト

閉会

運営メンバー (50音順，敬称略)

- 東貴範
- 阿部妙子
- 荒井ひろみ
- 井口誠
- 伊藤聡志
- 江口弘人
- 小栗秀暢
- 菊池浩明
- 黒政敦史
- 千田浩司
- 中川裕志
- 中村優一
- 西山賢志郎
- 野島良
- 波多野卓磨
- 濱田浩気
- 藤崎千尋
- 古川諒
- 馬瑞強
- 前田若菜
- 三浦尭之
- 村上隆夫
- 山田明
- 渡辺知恵美

各種担当

- システム
  - 井口さん，中村さん，西山さん，波多野さん，山田先生
- 事務局
  - 波多野さん，江口さん，東さん，山田先生，黒政さん
- 広報
  - 藤崎さん，中村さん，波多野さん
- 現地運営
  - 藤崎さん，中村さん，黒政さん
- 告知ポスター
  - 江口さん，千田先生，中村さん
- 表彰・CSS連携
  - 藤崎さん，山田先生，東さん

おわりに

- PWS Cup 参加者のみなさま,
  コンテストにご参加いただき,
  とっておきの情報を共有くださり,
  ありがとうございました.

- PWS Cup に参加されなかったみなさまも,
  議論を盛り上げていただき,
  ありがとうございました.

- 次回も(は)ぜひ PWS Cup にご参加ください

記念撮影

- 写ってもいいよという方は前方にお集まりください
- Webページや報告記事等に掲載する場合がございますので，ご了承ください

ご連絡

- 参加チームの皆様は，
  ポスターを忘れずにお持ち帰りください

PWS Cup セッション

休憩中

- 次は 11:30 から「ポスター発表前半」

PWS Cup セッション

ポスター発表前半

- 12:20まで．ホワイエで実施中
- チームIDが奇数のチームは必ずご発表ください

チームID	チーム名
01	堅あげポテト
03	ぼっちくんち
05	F.SE
07	Guardians of Privacy vol.23
09	みずほに恩返し
11	NYCU Quester
13	TCFSH
15	鋼鉄の錬金術師

昼休憩中

- 次は 13:40 から「ポスター発表後半」

ポスター発表後半

- 14:30まで．ホワイエで実施中
- チームIDが偶数のチームは必ずご発表ください

チームID	チーム名
02	神馬大 νττ
04	宮地研.exe
06	GOUTAI
08	ミズホノデシ
10	ステテコ清水
12	タケタケ
14	Enthusiasm

PWS Cup セッション

休憩中

- 次は 14:40 から「閉会」