femto-st SCIENCES & TECHNOLOGIES

Assainissements pour des analyses respectueuses de la vie privée en santé

J.-F. Couchot

FEMTO-ST/DISC/DEODIS

Séminaire RDI BMB

cnrs UBFC UNIVERSITÉ BOURGOGNE-FRANCHE-COMTÉ UNIVERSITÉ DE FRANCHE-COMTÉ SUP MICRO TECH ENSMM utbm université de technologie Belfort-Montbéliard

Anonymisation/assainissement : cadre légal

Quelques réglementations

- Décl°. univ. des droits de l'homme [1] : pas d'immixtion dans la vie privée
- Règles européennes sur l'IA. [2] : décisions algorithmiques critiques par IA. uniquement si explicables, sûres
  - ⇝ évaluation sur des données réalistes
  - ⇝ modèles et sorties : fuites maîtrisées d'information
- RGPD [3] : cadre protecteur par rapport aux données
  - ⇝ contraintes réduites sur les données anonymes
- e-privacy [4] : traitᵗ. des données personnelles par les opér. téléphoniques
  - ⇝ doit être fait à la volée (sans mémorisation)

Problématique d'anonymisation/assainissement de données

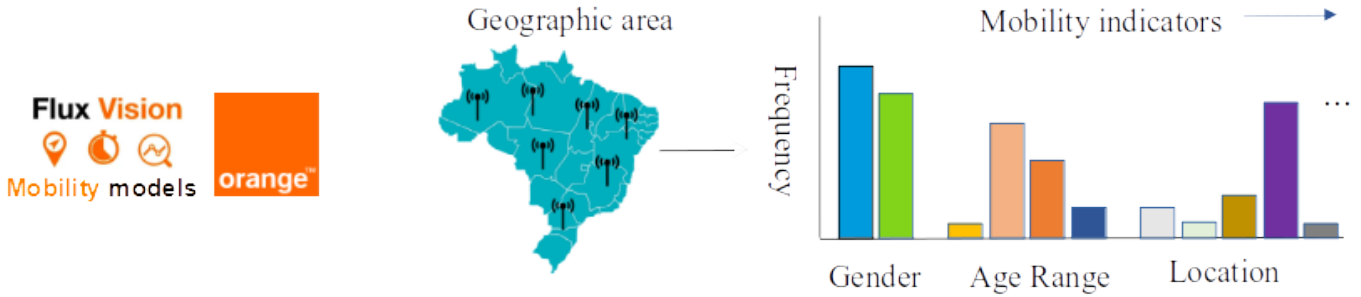
- Démarche pour des analyses respectueuses réglementairement
- Objectif : pour un niveau de protection défini, maximisation de l'utilité

1. https://www.un.org/fr/universal-declaration-human-rights/
2. https://www.europarl.europa.eu/news/fr/press-room/20230609IPR96212/les-deputes-sont-prets-a-negocier-les-regles-pour-une-ia-sure-et-transparente
3. https://www.cnil.fr/fr/lanonymisation-de-donnees-personnelles
4. https://www.economie.gouv.fr/files/files/directions_services/cge/e-privacy.pdf

Anonymisation/assainissement : comment ?

Approches syntaxiques sur les données : k-anonymat [5], dissociation [6]

- Données groupées par classes d'effectif $\geq k$
- Mise en place aisée (mais attaquable par connaissances supplémentaires)



Propriété probabiliste sur l'algorithme $\mathcal{M}$: ε-confidentialité différentielle (ε-DP) [7]

$$\forall D_1, D_2 \text{ (bases voisines)}, D, O \text{ (image)}, \frac{\Pr(D = D_1 | \mathcal{M}(D) = O)}{\Pr(D = D_2 | \mathcal{M}(D) = O)} \leq e^{\varepsilon} \frac{\Pr(D = D_1)}{\Pr(D = D_2)}$$

- Publicat°. de $\mathcal{M}(D) = O$: capacité à distinguer D_1 de $D_2 \approx$ inchangée
- Pratiq^{t}. : créat°. de mécanismes $\mathcal{M}$ aléatoires ajoutant un bruit contrôlé

---

5. Sweeney, L. (2002). k-anonymity : A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05), 557-570.

6. Terrovitis, M., Liagouris, J., Mamoulis, N., & Skiadopoulos, S. (2012). Privacy preservation by disassociation. arXiv preprint arXiv :1207.0135.

7. Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006, March). Calibrating noise to sensitivity in private data analysis. In Theory of cryptography conference (pp. 265-284). Springer, Berlin, Heidelberg.

Approches syntaxiques : contributions

Dissociation (découpage vertical & horizontal) : attaques

- Preuve d'existence d'attaques [8]
- Correctif [9] et preuves (correction, terminaison) de l'approche [10]

Réassociation (pour exploitation par ML) respectant une distribution [11]

- Métriques de Kulback-Leibler et Hellinger : pas adaptées aux ensembles
- Distance d'édition sur les arbres [12] combinée à une génération roulette

---

8. Barakat, Sara, Bechara Al Bouna, Mohamed Nassar, and Christophe Guyeux. "On the evaluation of the privacy breach in disassociated set-valued datasets." arXiv preprint arXiv :1611.08417 (2016).

9. AWAD, N., COUCHOT, J.-F., AL BOUNA, B., PHILIPPE, L. Ant-driven clustering for utility-aware disassociation of set-valued datasets. In : Proceedings of the 23rd International Database Applications & Engineering Symposium. 2019. p. 1-9.

10. AWAD, N., AL BOUNA, B., COUCHOT, J.-F., PHILIPPE, L. Safe disassociation of set-valued datasets. Journal of Intelligent Information Systems, 2019, vol. 53, no 3, p. 547-562.

11. AWAD, N., COUCHOT, J.-F., AL BOUNA, B., PHILIPPE, L. Publishing Anonymized Set-Valued Data via Disassociation towards Analysis. Future Internet, 2020, vol. 12, no 4, p. 71.

12. Zhang, K., & Shasha, D. (1989). Simple fast algorithms for the editing distance between trees and related problems. SIAM journal on computing, 18(6), 1245-1262.

ε-DP : quoi, qui, comment ?

Données multidimensionnelles

- Construire des histogr. à p. de tuples de d attributs $\{A_1, \ldots, A_d\}$ discrets

RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response

Úlfar Erlingsson
Google, Inc.
ulfar@google.com

Vasyl Pihur
Google, Inc.
vpihur@google.com

Aleksandra Korolova
University of Southern California
korolova@usc.edu

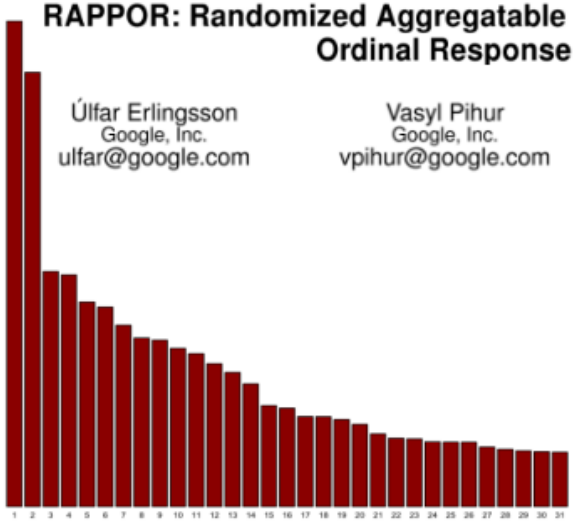


Figure 6: Relative frequencies of the top 31 unexpected Chrome homepage domains found by analyzing ∼14 million RAPPOR reports, excluding expected domains (the homepage "google.com", etc.).

Learning with Privacy at Scale

Differential Privacy Team, Apple

The Count Mean Sketch technique allows Apple to determine the most popular emoji to help design better ways to find and use our favorite emoji. The top emoji for US English speakers contained some surprising favorites.

Collecting Telemetry Data Privately

Bolin Ding, Janardhan Kulkarni, Sergey Yekhanin
Microsoft Research
{bolind, jakul, yekhanin}@microsoft.com

Windows Insiders in Windows 10 Fall Creators Update to protect users' privacy while collecting application usage statistics.

ε-DP : quoi, qui, comment ?

Données multidimensionnelles

- Construire des histogr. à p. de tuples de d attributs $\{A_1, \ldots, A_d\}$ discrets

Mécanismes $\mathcal{M}$ aléatoires pour 1 seul attribut à valeur dans $\{v_1, \ldots, v_k\}$

- $\mathcal{M}_{GRR}$ [13], $\Pr[\mathcal{M}_{GRR}(x) = v_i] = \begin{cases} p = \dfrac{e^\varepsilon}{k-1+e^\varepsilon} & \text{si } v_i = x \\ q = \dfrac{1}{k-1+e^\varepsilon} & \text{sinon} \end{cases}$
- $\mathcal{M}_{SUE}$ [14], v_i cod. bin., $\Pr[\mathcal{M}_{SUE}(v_i) = 1] = \begin{cases} p = \dfrac{e^{\varepsilon/2}}{e^{\varepsilon/2}+1} & \text{si } v_i = 1 \\ q = \dfrac{1}{e^{\varepsilon/2}+1} & \text{si } v_i = 0 \end{cases}$

Estimation de la fréquence d'apparition de v_i

- N pers., f_i (resp. r_i) la fréq. init. de v_i (resp. après application de $\mathcal{M}$)
- Estimateur $\hat{f}_i = \dfrac{r_i - q}{p - q}$, $\mathrm{Var}[\hat{f}_i] = \dfrac{q(1-q)}{N(p-q)^2} + \dfrac{f_i(1-p-q)}{N(p-q)}$

13. Kairouz, P., Bonawitz, K., & Ramage, D. (2016). Discrete distribution estimation under local privacy. arXiv preprint arXiv :1602.07387.

14. Erlingsson, Ú., Pihur, V., & Korolova, A. (2014, November). Rappor : Randomized aggregatable privacy-preserving ordinal response. In Proceedings of the 2014 ACM SIGSAC conference on computer and communications security (pp. 1054-1067).

ε-DP : contributions

Choisir (à la volée) le mécanisme minimisant la dispersion [15]

- ε, k : connus $\rightsquigarrow$ comparaison possible à priori des variances $\rightsquigarrow$ choix du mécanisme la minimisant

Conséquences de $\mathrm{Var}(\varepsilon/d) \geq \mathrm{Var}(\varepsilon)$ pour d attributs de sensibilité $\neq$

- Partage de $(j, \mathcal{M}(v_j))$, $j \in rand(1, d)$: équitable ?
- Proposition : 1 attribut choisi aléatoirement et assaini, les autres fictifs selon une distribution choisie [16] [17], estimateurs, variances et codes [18]

---

15. Arcolezi, H. H., Couchot, J. F., Al Bouna, B., & Xiao, X. (2022). Improving the utility of locally differentially private protocols for longitudinal and multidimensional frequency estimates. Digital Communications and Networks.

16. Arcolezi, H. H., Couchot, J. F., Al Bouna, B., & Xiao, X. (2021, October). Random sampling plus fake data : Multidimensional frequency estimates with local differential privacy. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management (pp. 47-57).

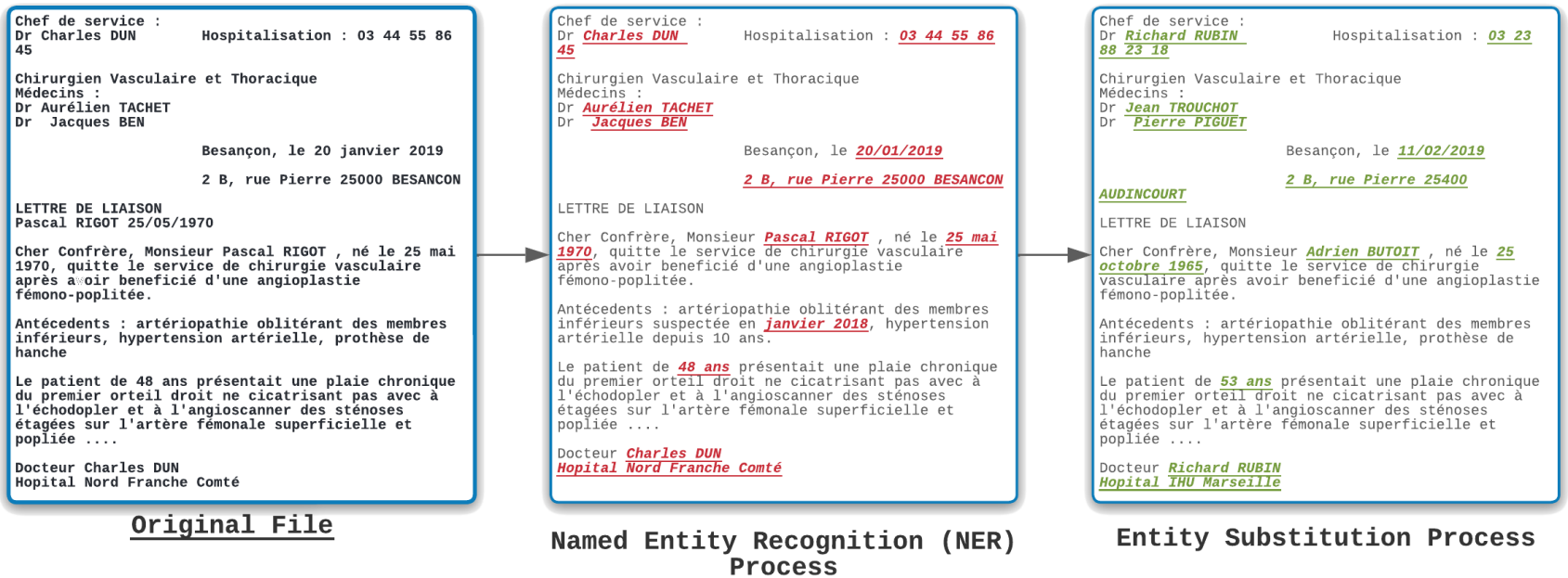
17. Arcolezi,H. H. , Gambs, S., Couchot, J.-F., & Palamidessi, C. : On the Risks of Collecting Multidimensional Data Under Local Differential Privacy. Proc. VLDB Endow. 16(5) : 1126-1139 (2023)

18. `https://hharcolezi.github.io/`

ε-DP : dé-identification de notes médicales

Démarche en 2 étapes

1. Détection des informations (quasi) identifiantes (efficacité NER)
2. *Assainissement* (optimisation de l'utilité médicale pour une fuite acceptée)



Détection d'entités

Apprentissages itératifs sur des datasets de l'HNFC

- de + en + gros, déidentifiés de + en + surement
- prélabélisés automatiquement et validés manuellement
- Modèle : hybrides, puis DL uniquement

Résultats de détection [19] [20] Dernoncourt [21].

Method	CamemBERT-ner			MEDINA			FlauBERT-ner			**Hybride**[18]			**Healthinf**[19]			*Dernoncourt*[20]		
Dataset	*HNFC*															i2b2		
Metric	*P*	*R*	F_1	*P*	*R*	F_1	*P*	*R*	F_1	*P*	*R*	F_1	*P*	*R*	F_1	*P*	*R*	F_1
PER	89	99	93.8	***98.2***	97.7	*98.2*	91.8	97.6	94.6	96.3	***99.8***	98	97.2	98.9	98	**98.2**	99.1	**98.6**
ORG	7.	21.8	11.1	32.6	24.8	28.1	16.9	34.1	22.6	41.1	*57.3*	47.8	*90*	51	*65.6*	**92.9**	**71.4**	**80.7**
LOC	46	67.2	54.6	98.8	81.1	89.1	75.7	66.3	70.7	88.4	***95.8***	92	***99.4***	94.4	***96.9***	95.9	95.7	95.8
DATE	NA			97.7	86.6	91.9	NA			97.7	86.7	91.9	***99.2***	*95.7*	*97.4*	99	**99.5**	**99.2**
AGE	NA			91.5	66.9	77.3	NA			91.5	66.9	77.3	*98.2*	*91.8*	*95*	**98.9**	**97.6**	**98.2**
TEL	NA			99.5	97.9	98.7	NA			***99.5***	97.9	98.7	99.4	***99.8***	***99.6***	98.7	99.7	99.2
REF	NA			NA			NA			NA			96.1	79.5	87	NA		
QID	NA			NA			NA			NA			77.2	32	45.3	**99.2**	**98.7**	**99**
Mic.-avg.	70.8	51.5	59.6	98.2	91.2	94.5	85.8	86.7	86.3	94.6	94.9	94.7	***98.5***	*96.4*	*97.4*	98.3	**98.5**	**98.4**

19. Y. Tchouka, J.-F. Couchot, M. C., D. Laiymani, P. Selles, A. Rahmani, C. Guyeux : De-Identification of French Unstructured Clinical Notes for Machine Learning Tasks. CoRR abs/2209.09631 (2022)

20. Tchouka, Y., Couchot, J.-F., Laiymani, D. : An Easy-to-Use and Robust Approach for the Differentially Private De-Identification of Clinical Textual Documents. HEALTHINF 2023 : 94-104

21. Dernoncourt, F., Lee, J. Y., Uzuner, O., & Szolovits, P. (2017). De-identification of patient notes with recurrent neural networks. Journal of the American Medical Informatics Association, 24(3), 596-606.

Assainissement

Utilité de l'ε-DP pour certaines entités ?

- $\forall o, x_1, x_2 \Pr(\mathcal{M}(x_1) = o) \leq e^{\epsilon} \Pr(\mathcal{M}(x_2) = o)$
- Probablement assainis avec la même valeur :
  - 08/01/42 et 14/03/18 (St. Hawking) :
  - Dijon et Beze (en BFC mais épidémiologiquement $\neq$)

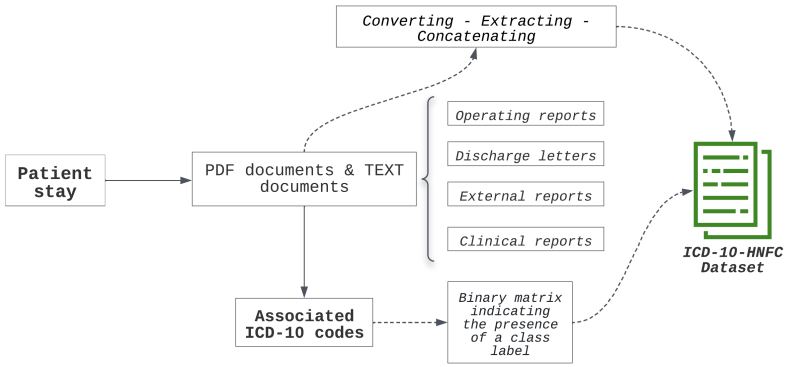
Assainissement intégrant de la $\varepsilon - d$ privacy [22]

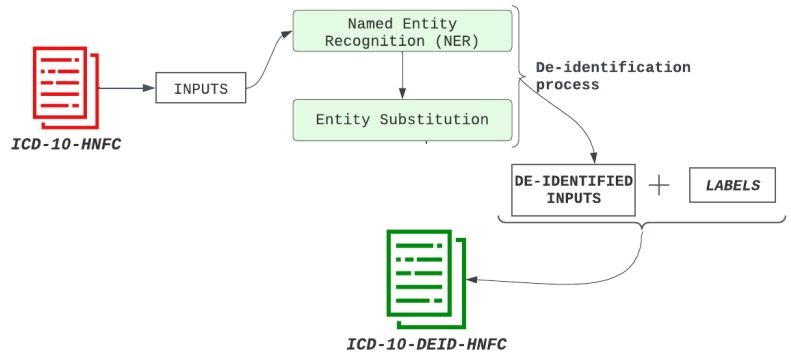
- Théorie : $\forall x_1, x_2, \Pr(\mathcal{M}(x_1) = o) \leq e^{d(x_1,x_2)\epsilon} \Pr(\mathcal{M}(x_2) = o)$
- Dates : $\mathcal{M}_{date}(x) = x + v$ t.q. $v \sim Lap(\frac{1}{\epsilon})$
- Localisations : $\Pr(\mathcal{M}_{loc}(x) = o)$ prop. à $e^{\varepsilon . d(x,o)}$ avec d distance en fonction de caractères épidémiologiques.

22. Chatzikokolakis, K., et al. "Broadening the scope of differential privacy using metrics." International Symposium on Privacy Enhancing Technologies Symposium. 2013.

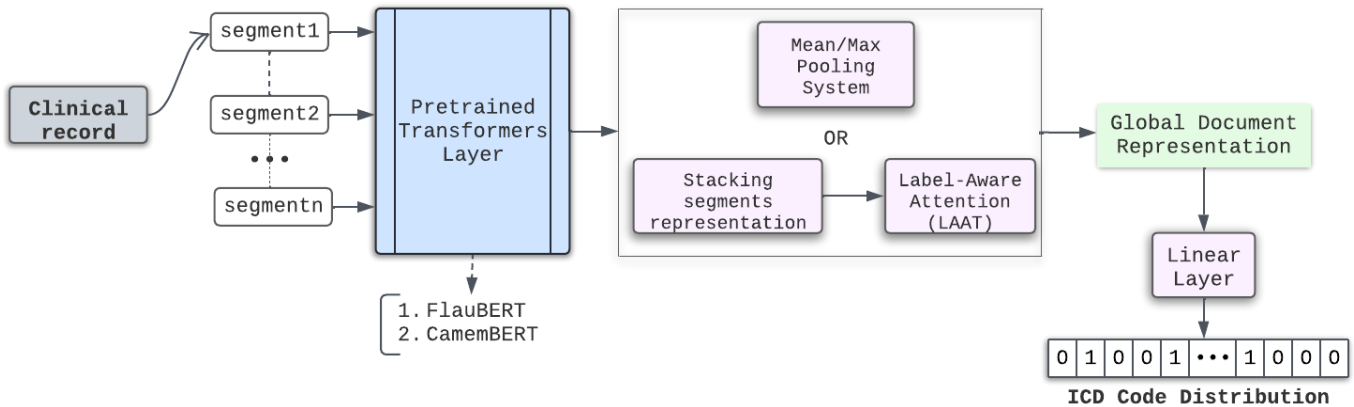
Association de codes CIM-10 [23]–1

Jeux de données





Architecture du modèle d'association de code CIM-10



23. Tchouka, Y., Couchot, J.-F., Laiymani, D., Selles, P., Rahmani, A. : Automatic ICD-10 Code Association : A Challenging Task on French Clinical Texts. CBMS 2023 : 91-96

Association de codes CIM-10–2

Résultats d'association p.r. à l'état de l'art

Models	Language	Dataset	Labels	F_1-score
PLM-ICD[24]	*English*	*MIMIC 2* *MIMIC 3*	*5,031* *8,922*	*0.5* ***0.59***
Dalloux[25]	*French*	*Personnel*	*6,116* *1,549*	*0.39* *0.52*
PROPOSAL	French	ICD-10-HNFC	6,160 1,564	**0.45** **0.55**
Dalloux[25]			6,160 1,564	0.27 0.35

Résultats sur l'utilité de la dé-identification

Dataset	Labels	**Precision**	**Recall**	F_1-**score**
ICD-10-HNFC	**6160**	**0.47**	**0.46**	**0.47**
ICD-10-DEID-HNFC		0.44	0.43	0.44
ICD-10-TAG-HNFC		0.43	0.41	0.42

24. Huang, C. W., Tsai, S. C., & Chen, Y. N. (2022). PLM-ICD : automatic ICD coding with pretrained language models. arXiv preprint arXiv :2207.05289.

25. BOUZILLE, G., & GRABAR, N. (2020). Supervised learning for the ICD-10 coding of French clinical narratives. Digital Personalized Health and Medicine : Proceedings of MIE 2020, 270, 427.

Visibilité de la thématique de dé-identification

Partenaires académiques

- Leader du projet ANR DIFPRIPOS 23-27 : intégrer de la DP dans PostgreSQL
  - Equipe INRIA COMETE/LIX polytechnique, LIFO, INSA Lyon
- UQAM (Montréal), NUS (Singapour), UA (Liban)

Partenaires non académiques

- Orange Flux vision (thèses CIFRE 20-23, et 23-26)
- HNFC (dé-identification de notes médicales)
- DALIBO (Société française de déploiement de PostsgreSQL)

Organisations et invitations récentes

- juin 23 : organisation de l'Atelier de Protection de la Vie Privée (APVP) du GT-PVP du GDR-Sécu
- juin 23 : invitation à présenter un exposé aux journées du GDR-Sécu

En chiffres

- 5 doctorant.e.s/docteur.e.s, 15 articles