

Une démarche méthodologique pour l'anonymisation des données personnelles sensibles



Anas ABOU EL KALAM (ENSIB - LIFO)

Yves DESWARTE (LAAS - CNRS)

Gilles TROUESSIN (Ernst & Young Audit)

Emmanuel CORDONNIER (ETIAM)

Plan

- ❖ Introduction
- ❖ Démarche d'analyse
- ❖ Exemples de scénarios
- ❖ Nouvelle solution
- ❖ Conclusions

Plan

❖ Introduction

Sécurité vs vie privée
Réglementation
point de départ ?

- ❖ Démarche d'analyse
- ❖ Exemples de scénarios
- ❖ Nouvelle solution
- ❖ Conclusions

"Privacy" : réglementation

❖ Internationale :

- A/RES/45/95 pour la réglementation des fichiers personnels informatisés

❖ Européenne :

- Directive 95/46/CE pour la protection des **données à caractère personnel**
- Directive 2002/58/EC concernant e le commerce électronique
- Résolution 98/1165 du 26 juin 1998 sur le droit au respect de la vie privée

❖ Nationale :

- Loi 78-17 "Informatique, fichiers et libertés", protection des **données nominatives**
- Loi 2002-303 du 4 mars 2002 relative aux droits des malades ; 94-458 pour la recherche médicale
- Code pénal, art. 226-13 (secret professionnel) et 15 (secret des correspondances) concernant les infractions pénales contre la vie privée

❖ Américaine :

- Loi fédérale "Privacy Act" de 1974
- "Privacy Protection act" (COPPA) adopté par le congrès en octobre 1998

Point de départ

- ❖ Les applications émergentes utilisent des données **personnelles**
- ❖ La législation existe

MAIS ...

- ❖ Pas le support technologique !

Par quoi commencer ?



- ⇒ **Comprendre le domaine**
- ⇒ **Démarche méthodologique**

Plan

- ❖ Introduction
- ❖ Démarche d'analyse

Définitions

Besoins & Objectifs & Exigences

Choix de solutions

- ❖ Exemples de scénarios
- ❖ Nouvelle solution
- ❖ Conclusions

Définitions [ISO 15408]

❖ Anonymat :

Impossibilité de déterminer le véritable **nom** de l'utilisateur
≡ ¬ (révélation de l'identité)

❖ "Pseunonymat" :

Anonymat + responsabilité

❖ Non-"chainabilité" :

Impossibilité d'établir un **lien** entre différentes **opérations** faites par un **même** utilisateur

❖ Non-observabilité :

Impossibilité de déterminer si une **opération** est en cours

Démarche

❖ Besoins :

- D'ordre général ; spécifiques à l'application, ...
- Ex : que veut on protéger ? Contre qui / quoi ?

❖ Objectifs :

- *Réversibilité* → chiffrement
- *Irreversibilité* → hachage
- *Inversibilité* : pseudonymisation → chiffrement à clés publiques
 - désanonymisation ⇒ procédure exceptionnelle

❖ Exigences :

- Comment exprimer les besoins de manière plus précise en tenant compte des attaques, de l'environnement, etc.
 - *Chaînage* :
 - *temporel* (toujours, parfois, jamais)
 - *spatial* (international, national, régional, local)
 - *spatio-temporel* (toujours et partout, parfois et partout, local et jamais)
 - *Robustesse* : à la réversion, à l'inférence

Démarche (suite)

❖ Choix de solutions :

▪ **Type**

- organisationnelle, contrôle d'accès
- mécanisme cryptographique
- fonction à sens unique

▪ **Pluralité**

- mono-anonymisation
- bi-anonymisation
- multi-anonymisation

▪ **Interopérabilité**

- transcodage (manuel)
- translation (mathématique)
- transformation (automatique)

Plan

- ❖ Introduction
- ❖ Démarche d'analyse
- ❖ Exemples de scénarios

Stockage et transfert de données médicales
Études épidémiologiques focalisées
Maladies à déclaration obligatoire

- ❖ Nouvelle solution
- ❖ Conclusions

Exemple de scénarios

❖ Stockage et transfert de données médicales :

- **Objectif** : réversibilité
- **Exigence** : robustesse à la réversion

❖ Maladies à déclaration obligatoire :

- **Besoins** : prévention, veille sanitaire, analyses épidémiologiques, ...
- **Objectif** : Anonymisation irréversible
- **Exigence** : chaînage universel, robustesse à la réversion et aux inférences, ...

❖ Études épidémiologiques focalisées :

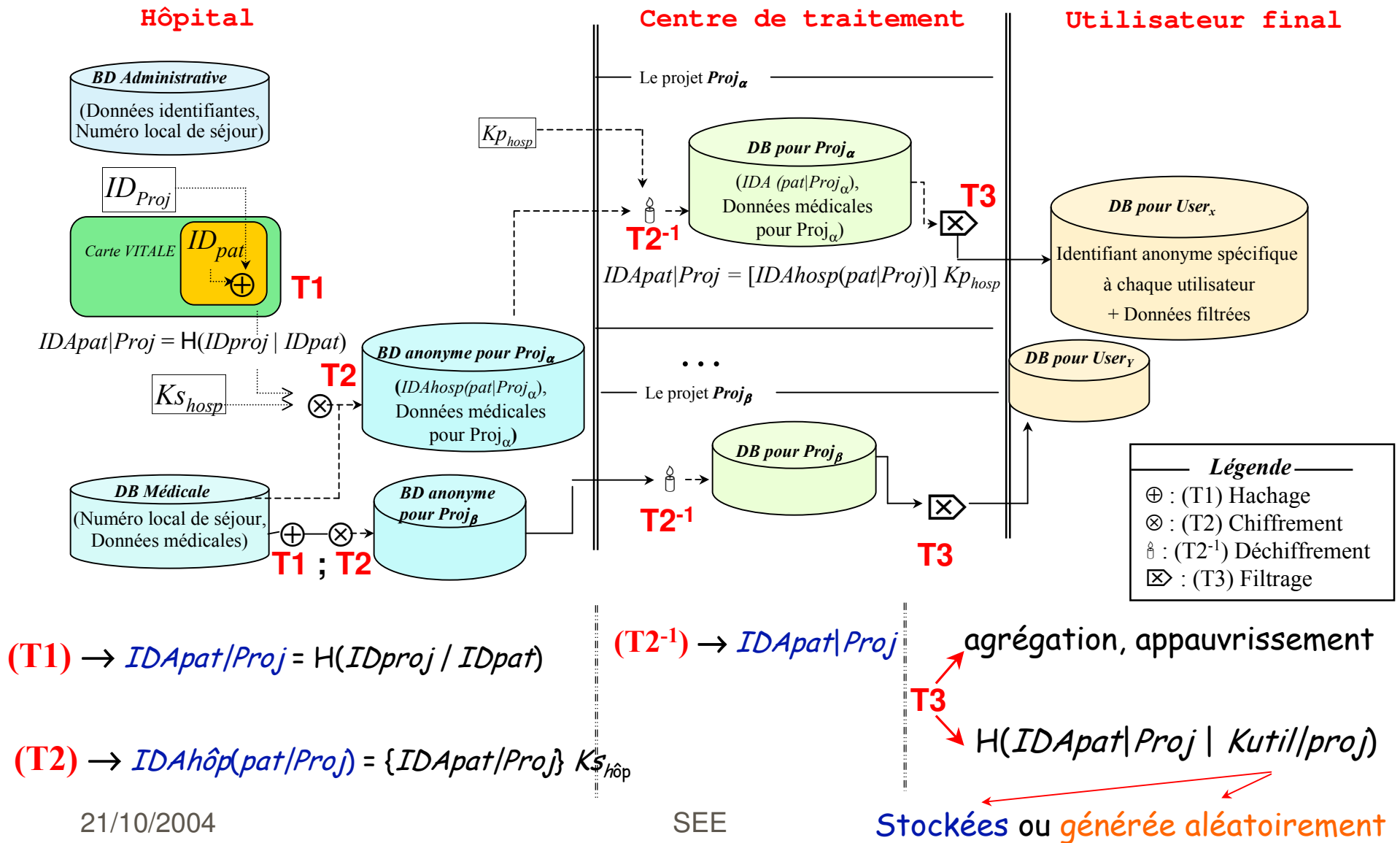
- **Besoins** : cacher les identités tout en ayant la possibilité d'identifier et d'informer les patients afin d'améliorer la qualité des soins
- **Objectif** : Anonymisation inversible (pseudonymisation)
- **Exigence** :
 - robustesse aux attaques par inférence
 - Type d'utilisation / d'utilisateur ⇒ Type du chaînage (temporel & géographique)

Plan

- ❖ Introduction
- ❖ Démarche d'analyse
- ❖ Exemple de scénarios
 - ❖ Nouvelle solution
- ❖ Conclusions

Schéma général
Discussion

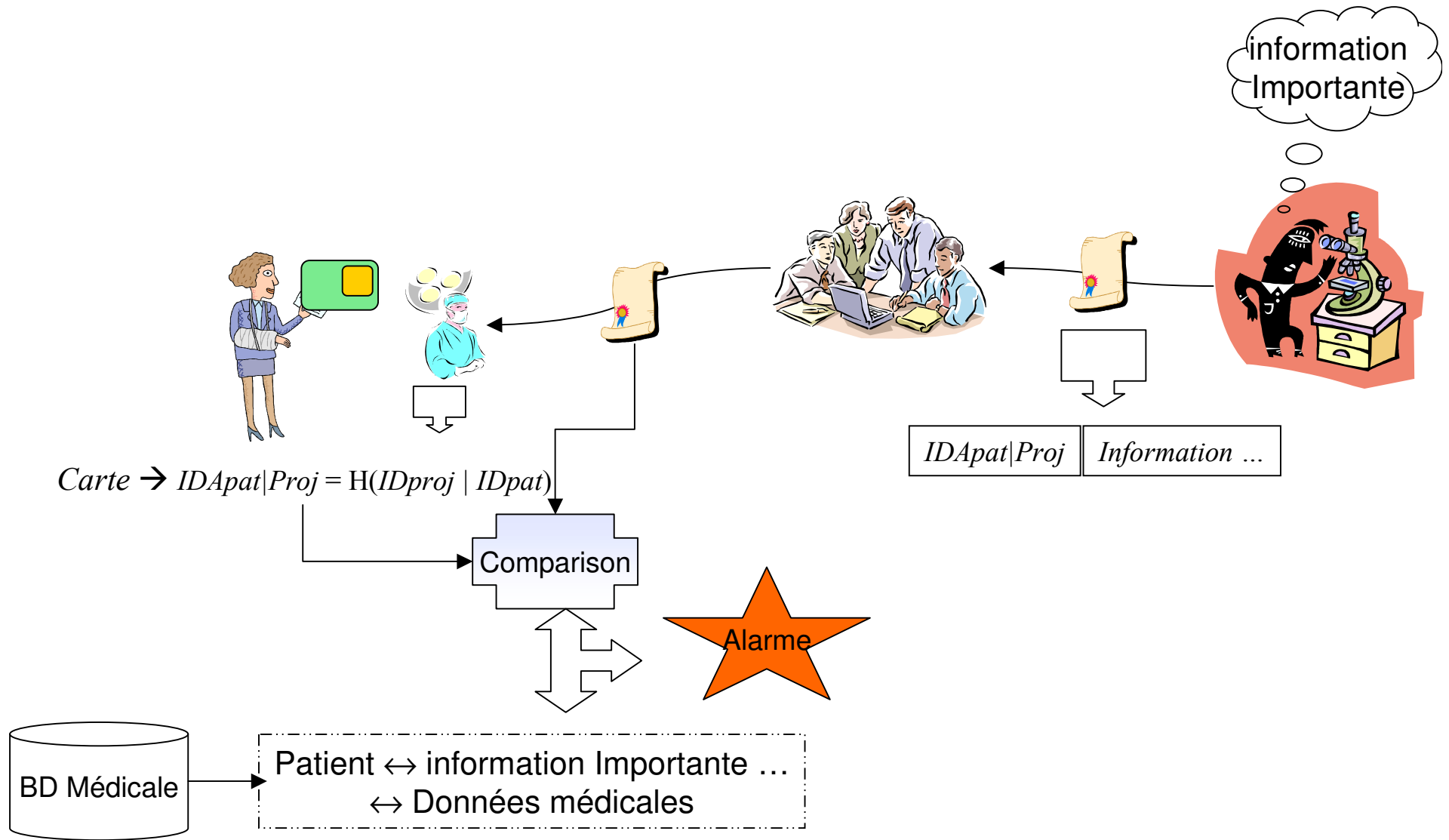
Nouvelle solution : Schéma général



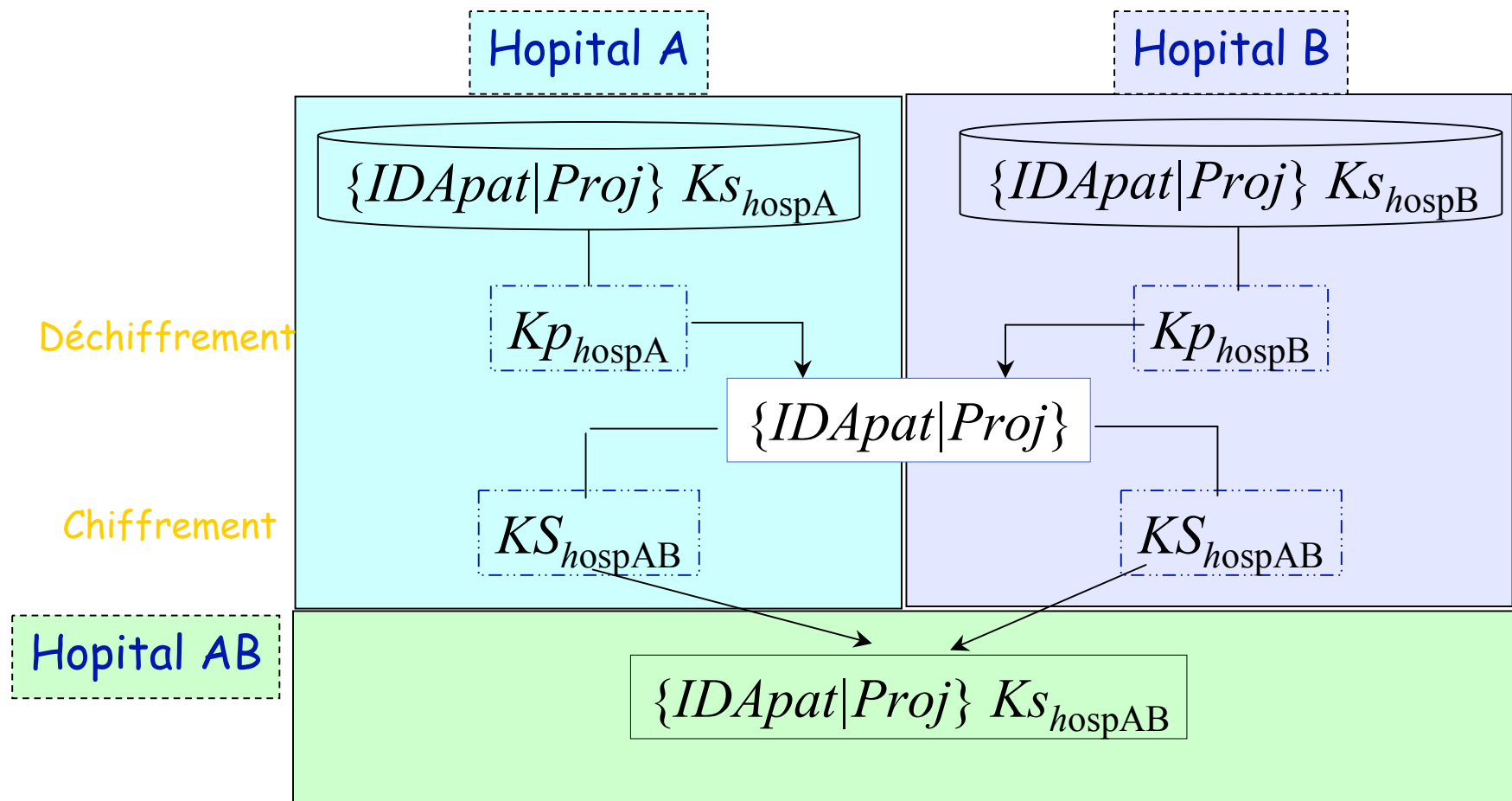
Nouvelle solution : discussion

- **Protection de l'identifiant anonyme du patient :**
 - *IDpat* est générée aléatoirement au sein de la carte
 - La carte est supposé suffisamment inviolable « tamper-resistant »
 - Le calcul $H(IDproj|IDpat)$ est effectué au sein de la carte
- **Absence de secret critique pour toute la population**
 - L'identifiant anonyme ne dépend que du (patient, projet)
 - Les identifiants sont situés dans des endroits différents
 - Les clés sont détenues par des personnes différentes
- **Consentement explicite du patient**
 - Lors de toute utilisation non-obligatoire, mais souhaitable, de ses données
 - Pour lever l'anonymat
- **Respect de la réglementation européenne / internationale**
 - Principe du moindre privilège
 - Finalité du traitement (objectif de l'utilisation)
 - Résistance aux attaques par dictionnaire, aux inférences par inversion, ...
- **Flexibilité**
 - Fusion de plusieurs établissements
 - Changement des objectifs de protection

Discussion : comment revenir aux identités ?



Discussion : fusion de plusieurs hôpitaux



Plan

- ❖ Introduction
- ❖ Démarche d'analyse
- ❖ Exemples de scénarios
- ❖ Nouvelle solution
- ❖ **Conclusions**

Conclusion

❖ Le risque "0" n'existe pas !

- **Démarche analytique**
 - Besoins → Objectifs → Exigences
- **Anonymisations**
 - Thématiques / en cascades

- ✓ Type de réversibilité ?
- ✓ Type de chaînage ?
- ✓ Forme de chaînage ?
- ✓ Robustesse à la réversion ?
- ✓ Robustesse à l'inférence ?

Mais aussi

- **Solutions organisationnelles**
 - Analyse des risques d'abduction
 - Politique de contrôle d'accès
- **Solutions techniques**
 - Architecture adaptée
 - Mécanismes appropriés (de contrôle d'accès)
 - IDS
 - Brouillage
 - Filtrage

Merci



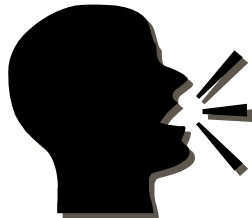
Questions ?



Remarques ?



Commentaires ?



Avis ?