# Définir le risque associé à un jeu de données

## Plan de gestion des données

Éric Quinton

eric.quinton@irstea.fr

IRSTEA - *RSSI*

ANF RENATIS

# Sommaire

# Que représente une donnée ?



Une donnée n'a aucune valeur intrinsèque :

- elle dépend de son contexte d'acquisition
  - précision de la mesure
  - protocole...
- ou de son traitement
  - analyse statistique
  - synthèse...
- c'est une représentation de la réalité

# Que représente une donnée ?





Une donnée n'a aucune valeur intrinsèque :

- elle dépend de son contexte d'acquisition
  - précision de la mesure
  - protocole...
- ou de son traitement
  - analyse statistique
  - synthèse...
- c'est une représentation de la réalité

Elle acquiert de la valeur dès qu'elle est générée :

- coût d'acquisition
- plus-value dans la connaissance...

# Que représente une donnée ?





Une donnée n'a aucune valeur intrinsèque :

- elle dépend de son contexte d'acquisition
  - précision de la mesure
  - protocole...
- ou de son traitement
  - analyse statistique
  - synthèse...
- c'est une représentation de la réalité

Elle acquiert de la valeur dès qu'elle est générée :

- coût d'acquisition
- plus-value dans la connaissance...

Dans certains cas, elle doit être protégée

# Les menaces informatiques n'ont jamais été aussi nombreuses
# Quelques exemples parmi d'autres

02/01/2015 : Les données de deux millions d'abonnés du site de TF1 ont été piratées. Les hackers détiennent les RIB et autres informations sensibles de ces internautes.

05/05/2015 : Les états -Unis (Office of Personnal Management) victime de piratage. Plus de 4 millions de données personnelles de personnels fédéraux piratées ;

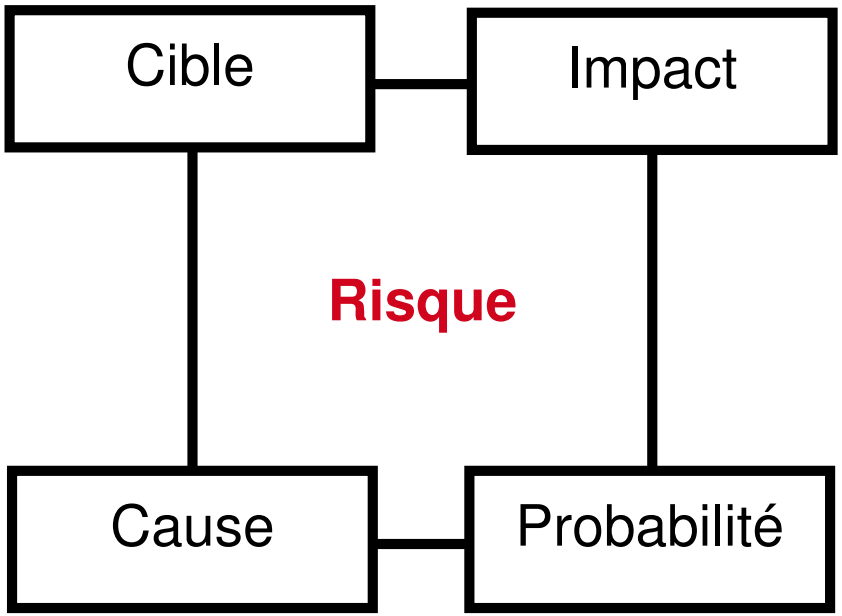
05/05/2015 : Arnaque aux faux virement : Vol de 15 millions d'euros à Intermarché

18/07/2015 : Piratage du site de rencontres adultères Ashley Madison

28/07/2015 : Les e-mails de hauts gradés de l'armée américaine piratés

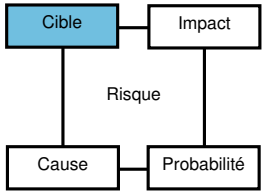
# Qu'est que le risque ?





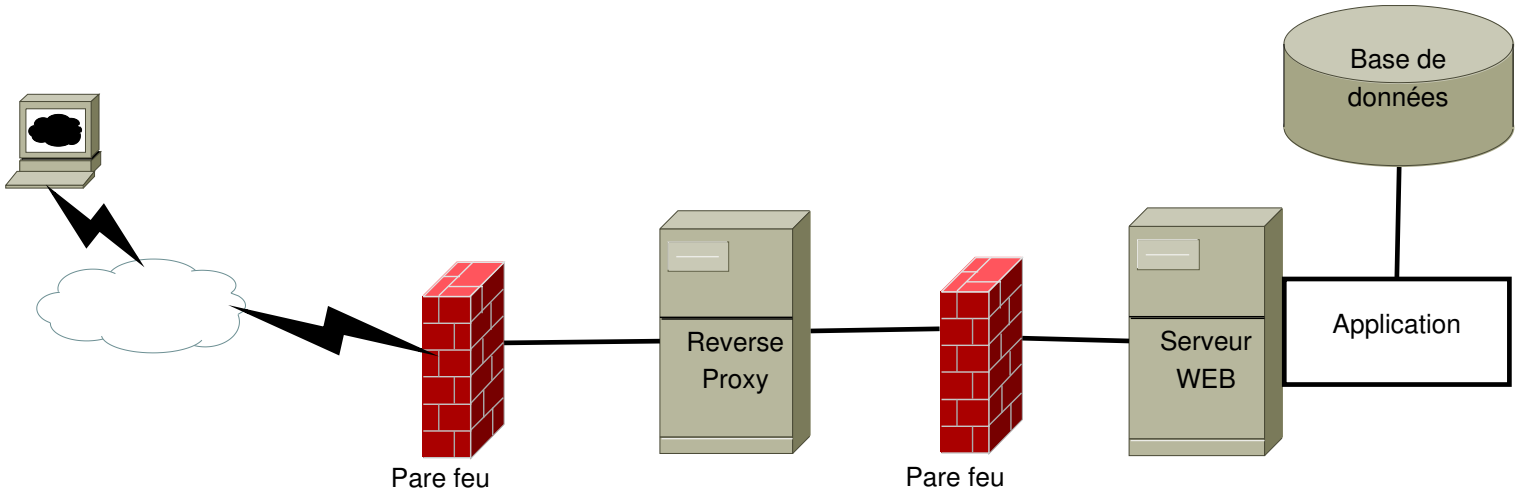

Le risque est souvent défini comme la conjonction entre :

- une cause / menace / événement ;
- une occurrence / probabilité / vraisemblance ;
- une cible ;
- un impact / gravité / conséquence.




# Définir la cible

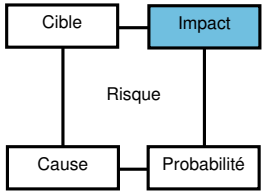
Un projet informatique s'inscrit dans une architecture :




Définir la cible : définir le périmètre de l'étude du risque.

Exemples : une **application couplée à une base de données**, ou un **jeu de données** => **PMD**

# Comment définir l'impact ou la gravité ?

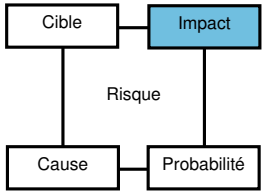





L'impact s'évalue selon trois critères (notés CID) :

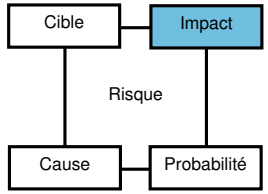
- la **confidentialité** : qui peut prendre connaissance des informations ou accéder au système mis en place ?
- l'**intégrité** : peut-on supporter la perte ou la corruption de données ?
- la **disponibilité** : le système peut-il être arrêté, et pendant combien de temps ?

Les critères sont classés selon des échelles (de 1 à 4 en général) définies par l'entreprise (exemples issus d'IRSTEA)




# La confidentialité

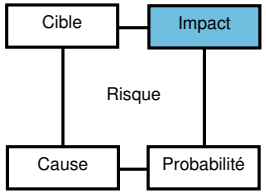
| Niveau | Nom | Contenu |
|---|---|---|
| 1 | public | Les informations traitées par le système sont accessibles à tous |
| 2 | limité | Les informations ne doivent être accessibles qu'au personnel et aux partenaires |
| 3 | réservé | Les informations ne doivent être accessibles qu'aux personnes impliquées dans le traitement des données |
| 4 | privé | Les informations ne doivent être accessibles qu'aux personnes identifiées et ayant besoin d'en connaître |




# L'intégrité

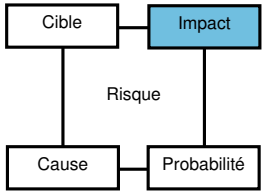
| Niveau | Nom | Contenu |
|---|---|---|
| 1 | Altérable | Les informations traitées peuvent ne pas être intègres |
| 2 | Détectable | L'information peut ne pas être intègre, si l'altération est identifiable et s'il est possible de revenir à une situation antérieure (sauvegarde) |
| 3 | Maîtrisé | L'information peut ne pas être intègre, si l'altération est identifiée et les informations récupérables (duplication en temps réel) |
| 4 | Intègre | L'information doit être rigoureusement intègre en toute circonstance, au besoin par un mécanisme de signature ou de chiffrement |

# La disponibilité




| Niveau | Nom | Contenu |
|---|---|---|
| 1 | Faible | L'application peut être indisponible pendant plus de 72 heures |
| 2 | Importante | L'application doit être disponible dans les 3 jours |
| 3 | Critique | L'application doit être disponible dans les 24 heures |
| 4 | Vitale | L'application doit être disponible dans les 4 heures |

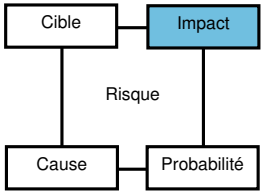
# L'estimation de l'impact en cas de défaillance




Pour chaque critère, 4 niveaux d'impact selon 4 thématiques :

| Nature | **1 limité** | **2 important** | **3 grave** | **4 critique** |
|---|---|---|---|---|
| **Fonctionnement interne** | perturbation limitée | Perturbation significative | Désorganisation très importante | Désorganisation durable |
| **Pertes financières** | non significatif | pertes < 50 K€ | < 200 K€ | > 200 K€ |
| **Responsabilité** | Plaintes d'usager pour un dysfonctionnement | Recours pouvant annuler une procédure | Plainte au civil | Plainte au pénal |
| **Atteinte à l'image** | impact limité | altération significative | altération très importante | altération définitive |

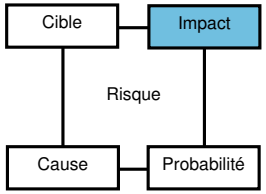
# L'étude débouche sur un tableau récapitulatif...






| | C | I | D |
|---|---|---|---|
| *Niveau à respecter* | | | |
| impact sur le fonctionnement interne | | | |
| impact financier | | | |
| impact en terme d'image | | | |
| impact en terme de responsabilité | | | |
| **impact maximal par critère** | | | |

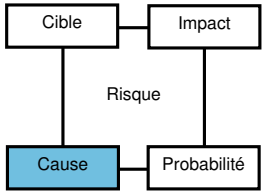
# ...qui est reportée dans le DMP






| Critère | Niveau de sécurité requis | Impact maximal en cas de défaut de sécurité |
|---|---|---|
| Confidentialité | | |
| Intégrité | | |
| Disponibilité | | |

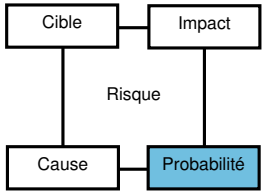
# Quelles menaces/causes à prendre en compte ?




- Les menaces sont très variées et évolutives ;
- des recueils de bonnes pratiques existent pour quasiment tous les domaines :
  - `http://www.ssi.gouv.fr/administration/bonnes-pratiques/` : site de l'ANSSI ;
- pour les applications WEB :
  - `http://www.ssi.gouv.fr/uploads/IMG/pdf/NP_Securite_Web_NoteTech.pdf`
  - `http://www.owasp.org` : *Open Web Application Security Project*, et le sous-projet ASVS (*Application Security Verification Standard Project*)

Il est plus simple de se référer à des recueils d'exigences que de chercher à inventorier toutes les menaces possibles.

# Comment estimer la probabilité ?




la probabilité d'occurrence d'une menace doit être évaluée en fonction du risque associé :

- plus les informations sont sensibles, plus le risque est grand d'être confronté à des attaques sophistiquées
- 3 niveaux de risque
  - attaque opportuniste (on teste au hasard)
  - attaque ciblée
  - attaque concertée avec travail de préparation important

Il est plus simple de considérer que tout événement peut se produire.

**Ne dites pas**

risque-t-on d'être attaqué ?
mais :
quand le sera-t-on ?

# En résumé



Au moment de l'élaboration d'un DMP :

- réalisez une analyse de risque
  - permet de déterminer les besoins en C, I, D
  - identifie les impacts en cas de problème
- intégrez les résultats dans le DMP
  - améliore la vision quant à l'exploitation possible des informations
  - sera utile au moment de la mise en place des processus d'acquisition ou de valorisation
- est souvent très rapide à faire (20' dans la plupart des cas)...
  - ... quand les formulaires adéquats sont disponibles !