

## Brevet fédéraux ICT

Elaborer des stratégies  
d'archivage, de sauvegarde  
et de restauration



# Sommaire

Introduction.....	7
<b>Chapitre 1</b>	
Analyse des besoins.....	9
<b>Chapitre 2</b>	
Stratégies de gestion de l'information.....	15
2.1 Identification des informations.....	16
2.2 Exigences d'archivage.....	21
2.3 Les solutions logicielles.....	27
<b>Chapitre 3</b>	
Classification pour la sauvegarde.....	39
3.1 Les enjeux.....	40
3.2 Classification métier.....	45
3.3 Classification technique.....	48
3.4 Objectifs de disponibilité.....	53
<b>Chapitre 4</b>	
Techniques et emplacements de sauvegarde.....	57
4.1 Emplacements de stockage des sauvegardes.....	59
4.2 Supports de sauvegarde.....	62
4.3 Types de sauvegarde.....	65

## **Chapitre 5**

Stratégie de sauvegarde.....	75
5.1 Collecte des informations nécessaires.....	78
5.2 Choix des mesures.....	79
5.3 Stratégie d'historisation.....	81
5.4 Planification.....	82
5.5 Responsabilités.....	83
5.6 Plan de restauration.....	84
Conclusion.....	87
Lexique.....	106
Table des illustrations.....	111
Table des matières.....	113



## Introduction

---

Si l'on voulait présenter le contexte en quelques mots, on pourrait simplement rappeler que la mission d'un système d'information consiste à soutenir les opérations d'une organisation – une entreprise ou une administration – et que celle de l'équipe IT consiste à soutenir le système d'information. Toutefois, lorsqu'on analyse davantage les activités du département informatique, on constate que c'est surtout le système, bien davantage que l'information, qui mobilise son temps, son budget et ses employés.

En un sens, le moyen occulte la fin : à force de se préoccuper de fournir toujours plus d'automatisation des traitements, de déployer des applications toujours plus sophistiquées, en assurant à la fois leur performance et leur sécurité, on finit par ne plus avoir le temps de s'intéresser à ce que transportent les « tuyaux » du système.

Il existe une autre raison, essentielle, à cette focalisation sur le système au détriment de l'information : les professionnels IT sont formés à concevoir et gérer des systèmes alors que les informations appartiennent aux métiers de l'organisation. Quant aux

utilisateurs, ignorant le plus souvent les possibilités du système, ils s'impliquent peu dans la gestion des informations et se contentent d'en produire encore et encore, seulement préoccupés de retrouver la bonne information au bon moment.

Les informations sont donc stockées dans le système sans réelle considération pour leur importance vis-à-vis des processus métier ni pour leur degré d'actualité. Ne parlons même pas de leur unicité : des versions de fichiers similaires voisinent sans que personne ne sache vraiment quelles versions sont caduques et les messageries regorgent de doublons parfaitement identiques.



### Analogie

*Lorsque nous faisons l'acquisition d'un nouvel objet, nous cherchons une nouvelle place pour lui dans la maison, mais comme beaucoup détestent jeter et préfèrent tout garder au motif que « ça peut servir », il faut donc constamment trouver de nouveaux rangements, tout comme la capacité de stockage d'un SI doit sans cesse croître.*

À la longue, le SI finit par ressembler aux caves ou aux galetas des habitations : on y stocke tout et on ne jette jamais rien, l'espace se réduit mais personne ne trouve le temps de faire le tri. Nous utiliserons d'ailleurs tout au long de ce manuel l'analogie entre la maison et le système d'information, car on y retrouve la même nette opposition entre une construction difficile à modifier et un contenu qui ne cesse de changer et d'augmenter.

Le premier chapitre sera consacré aux attributs qui caractérisent les informations et qui, correctement gérés, permettent d'effectuer un tri et une gestion différenciée de chaque type d'information.

Le second chapitre porte sur les stratégies de gestion des informations, à l'aide de leurs attributs et des outils logiciels appropriés.

Nous nous intéresserons ensuite à la sauvegarde, tout d'abord à ses critères et techniques, avant de terminer sur l'élaboration de l'indispensable stratégie de sauvegarde et de restauration.



# Chapitre 1

## Analyse des besoins

---

Comme toujours, la recherche d'une solution débute par l'analyse du besoin, comme le veulent les bonnes pratiques de la qualité. Car si les solutions sont souvent inadaptées, incomplètes voire totalement absentes, c'est précisément parce que cette analyse n'a pas été faite, ou seulement de manière implicite.

Pourtant, définir à quelle fréquence les informations sont accédées, quelle est leur durée de validité et combien de temps au maximum elles peuvent être indisponibles est fondamental pour que le SI lui-même réponde aux besoins.

L'analyse du besoin est d'autant plus nécessaire dans le cas de l'archivage et de la sauvegarde que les solutions techniques sont nombreuses et variées, en fonctionnalités comme en coûts. Pour chaque organisation, la solution idéale doit être recherchée : celle qui couvre les exigences vis-à-vis des différentes catégories d'information tout en respectant les contraintes liées au budget, à

l'intégration dans l'existant, à la charge de travail nécessaire pour la maintenance, etc.

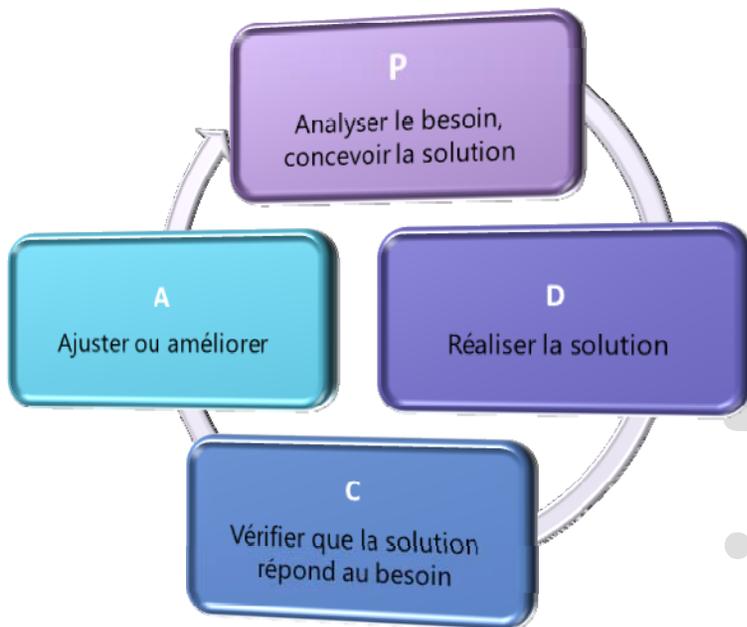


Figure 1 : la roue de Deming ou modèle PDCA<sup>1</sup>

<sup>1</sup> tous les termes soulignés en pointillés sont définis dans le lexique, page 106 et suivantes.

Pour chaque catégorie d'information, les exigences suivantes doivent être identifiées :

-  **disponibilité** : le terme recouvre en fait deux notions différentes, d'une part le temps que l'on peut attendre pour accéder à l'information (**RTO**), d'autre part la quantité de données que l'on peut accepter de perdre définitivement (**RPO**)
-  **intégrité** : l'information est aussi correcte et complète qu'à l'instant où elle a été introduite dans le SI, c'est-à-dire qu'elle n'a pas été modifiée ou tronquée par des opérations non souhaitées.
-  **traçabilité** : de plus en plus, les réglementations exigent que les entreprises puissent prouver l'intégrité de leurs données, ce qui implique que les modifications soient contrôlées et journalisées – ou interdites, auquel cas de nouvelles versions sont générées à chaque modification, laissant l'état précédent de l'information intact (historisation)
-  **confidentialité** : la majorité des informations qui sont produites dans une entreprise lui appartient strictement et ne peut être utilisée par d'autres sans son autorisation ; toute catégorie d'information doit donc être accompagnée de ses autorisations de diffusion<sup>1</sup>.



### Jargon

On appelle **RTO** (Recovery Time Objective) l'objectif de délai de reprise et **RPO** (Recovery Point Objectif) l'état du SI auquel la restauration doit permettre de revenir. Les deux objectifs s'expriment en heures.

---

<sup>1</sup> La catégorisation des niveaux de confidentialité, d'intégrité et de disponibilité est détaillée dans les ouvrages de votre cursus traitant spécifiquement de la sécurité de l'information. Elle est ici simplement résumée en annexe 1, page 91.

Concernant les contraintes, celle du budget est prédominante : stocker des fichiers et des bases de données a un coût, celui du système de stockage et de sa maintenance. Bien que le coût du gigaoctet stocké sur support magnétique ne cesse de diminuer, le volume de données croît de telle manière que le coût global du stockage continue d'augmenter.



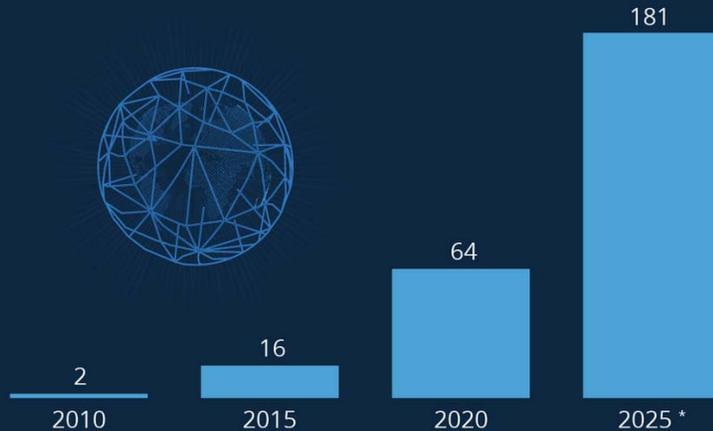
### Analogie

*Vous aimez tant les livres que vous en possédez des centaines, mais vous ne gardez au salon que ceux que vous souhaitez relire (disponibilité) et avez rangé les autres à la cave (archivé) dans des boîtes en plastique (intégrité). Une fois la cave pleine, vous devrez soit louer un espace de stockage, soit décider de jeter quelques boîtes...*

La virtualisation des machines et l'intégration au SI de sources d'information en provenance d'Internet n'ont clairement pas inversé cette tendance, bien au contraire.

## Le Big Bang du Big Data

Estimation du volume de données numériques créées ou répliquées par an dans le monde, en zettaoctets



Un zettaoctet équivaut à mille milliards de gigaoctets.

\* Prévvision en date de mars 2021.

Sources : IDC, Seagate, Statista



statista

Figure 2 : prévision de l'augmentation du besoin en stockage à l'horizon 2025

Il deviendra donc de plus en plus coûteux de maintenir la stratégie actuelle consistant à conserver la totalité de l'information constamment accessible au sein du système. Cela devient également de plus en plus risqué car les réglementations sont de plus en plus exigeantes sur le stockage de données personnelles. Ainsi, des fichiers qui ne sont plus utilisés depuis longtemps peuvent causer un dommage grave à l'organisation qui les stocke si des données personnelles sont volées.



### Jargon

**Big Data** : nom donné aux grands volumes de données générés par les réseaux de communication modernes (Internet, smartphones, puces RFID et capteurs, etc.)