Un service de protection des données professionnelles des personnels.

Une solution pour les composantes et les unités de recherche.

Cette note présente le principe, les objectifs et les avantages d'une solution de protection des données professionnelles pour les personnels de l'université et les personnels hébergés quel que soit le poste de travail sur lequel sont stockées ces données.

Un nouveau service de protection des données professionnelles des personnels ?

Une solution plus particulièrement adaptée aux besoins des composantes et des unités de recherche.

Résumé du projet.

Offrir aux personnels de l'université et aux personnels hébergés, travaillant dans les composantes et dans les unités de recherche, un service unique de sauvegarde automatisée et de restauration à la demande de leurs données présentes sur tous leurs postes de travail (ordinateurs fixes et portables, tablettes et smartphones). Ce service doit être accessible depuis n'importe quel point de connexion à Internet (*sauvegarder et restaurer de n'importe où*).

La DSI estime le coût logiciel d'un tel projet entre 21 000 €HT et 40 000 €HT par an pour 1000 utilisateurs.

Des données dispersées et mal protégées.

50% des sondés ont déjà perdu des données.

20% des sondés sauvent leurs données sur des *clouds* publics.

25% du temps de travail effectué en dehors des locaux de l'université.

Chaque personnel dispose en moyenne de 1,4 poste de travail.

Rappel des objectifs.

La sécurisation des données professionnelles des enseignants et des chercheurs est le principal objectif de ce projet. Suite à une enquête menée par la DSI auprès des utilisateurs potentiels de ce futur service, nous pouvons décrire les attentes auxquelles doit répondre ce nouveau service :

- Un seul service pour toutes les données, quel que soit le poste de travail de l'utilisateur où elles sont stockées ;
- L'utilisateur a une vue consolidée de l'ensemble de ses données professionnelles sauvegardées ;
- Les sauvegardes sont automatisées et après une configuration initiale, ne requièrent pas l'intervention de l'utilisateur ;
- Les opérations de restauration sont simples et peuvent être menées par un non informaticien ;
- Un service disponible en tous points de l'Internet ;
- Un service disponible sur Windows, sur Mac OS et sur Linux ;
- Un service géré centralement par la DSI, mais dont la gestion peut être en partie déléguée à un informaticien de composante ou de laboratoire ;
- Un service disponible pour tous les personnels d'une unité de recherche, quel que soit leur employeur.

Le contexte.

Les personnels de l'université travaillent sur différents types de poste de travail ; des ordinateurs fixes ou portables, des tablettes ou des smartphones sur lesquels sont conservées des données professionnelles. Il est souvent compliqué pour l'utilisateur de garantir la protection de ses données, surtout en cas de mobilité.

Lorsqu'il est sensibilisé à l'importance de la sauvegarde de ses données professionnelles, l'utilisateur met souvent en œuvre de multiples mécanismes et moyens de sauvegarde qui ne garantissent ni l'intégrité, ni la pérennité de ses données.

Pour pallier les risques réels de perte et de fuite de données, la mise en place d'un service centralisé de sauvegarde des postes de travail est à l'étude au sein de la DSI et en liaison avec plusieurs informaticiens de composantes ou de laboratoires.

Une enquête menée fin 2014 par la DSI auprès des personnels de l'université a permis de mettre en lumière plusieurs points.

Plusieurs postes de travail et dispersion des données.

Les utilisateurs n'ont rarement qu'un seul poste et travaillent souvent depuis plusieurs terminaux (en moyenne 1,4 équipements professionnels par utilisateur).

Il est rare qu'un utilisateur n'utilise qu'un seul poste fixe. La tendance est à l'utilisation d'un ordinateur portable en poste principal et souvent à au moins un autre poste (fixe professionnel, personnel, ou mobile). On observe alors une dissémination des données utilisateurs dont la protection intégrale devient difficile.

Les moyens informatiques individuels sont multi plates-formes[1] avec une diversité encore accrue liée à l'explosion des terminaux mobiles.

Mobilité et connexion au réseau.

Les utilisateurs sont rarement sédentaires. La moitié des utilisateurs estime être en dehors de l'université plus de 25% du temps, soit plus d'une journée par semaine en moyenne.

Aussi, si l'on souhaite la protection continue des données utilisateurs, le service devra être disponible et accessible de n'importe où via Internet.

Cette mobilité impacte les moyens d'accès au réseau. Les accès de type Wi-Fi sur site de l'université augmentent avec un tiers des utilisateurs qui l'utilisent davantage qu'une connexion filaire. Certains utilisateurs utilisent exclusivement le Wi-Fi, pour des raisons pratiques et techniques (terminaux sans accès filaire).

En mobilité à l'extérieur d'un site de l'Université, 80% des sondés expriment le besoin d'une connexion au réseau; ils utilisent alors des réseaux domestiques, d'entreprises ou publiques.

Les données utilisateurs sont alors hors de protection des services informatiques de l'université et exposées à de nombreux risques (écoutes, attaques informatiques, virus, etc.).

[1] Windows, Mac OS, Linux, iOS, Android.

Habitude de sauvegarde.

La moitié des sondés déclarent avoir déjà été confrontés à une perte de données.

Ces dommages sont, dans la majorité des cas, liés à des dysfonctionnements matériels mais également à des vols ou des pertes et parfois à des programmes informatiques malveillants.

Si une grande majorité des sondés sont sensibilisés aux questions de protection et de sauvegarde de leur données, deux tiers d'entre eux utilisent des moyens jugés non fiables (supports ou périphériques amovibles) et une part non négligeable (20% des sondés) déclarent externaliser leurs données professionnelles vers des solutions Cloud grand public.

L'utilisation de supports ou périphériques amovibles comme moyen de sauvegarde n'est pas recommandée ; en effet la fiabilité et la pérennité de ces moyens ne sont pas assurées. L'utilisateur conscient de ce problème pourra multiplier les supports (en double, en triple), mais également les manipulations et les coûts.

Les conditions générales d'utilisation des offres *cloud* public sont souvent méconnues des utilisateurs. Et pourtant, en leur confiant leurs données, même s'ils en restent généralement propriétaires, ils cèdent cependant des droits d'exploitation et de modification aux opérateurs de ces *clouds* publics.

Ceux-ci, principalement américains, sont également soumis au *Patriot Act*, qui permet aux autorités américaines d'accéder aux données de l'utilisateur sans son autorisation ou obligation de l'informer.

Les personnes ne faisant pas de sauvegarde de leurs données, expliquent ne pas le faire principalement pour des raisons pratiques (contraintes de temps) et techniques (compatibilité des systèmes et volumétries disponibles).

Mise en œuvre du projet.

L'enquête et la communication qui l'a accompagnée ont permis d'ouvrir le débat avec un certain nombre d'informaticiens et d'utilisateurs.

Un groupe de travail s'est mis en place ; ses membres se réunissent régulièrement et échangent au sujet des formes que pourraient prendre le service. Ce groupe est constitué d'Administrateur Système et Réseau (ASR) volontaires et intéressés par la mise à disposition d'un tel service auprès de leurs utilisateurs. Ces ASR sont affectés dans des composantes ou des unités de recherche.

Les DSI des cotutelles principales (INSERM, CNRS) ou leur représentant, ont été informés de la mise en place du projet. Les membres du groupe de travail représentent, dans la grande majorité des cas, des unités de recherche ou des départements qui n'ont pas de solutions de sauvegarde à proposer à leurs utilisateurs et qui n'ont pas pu mettre en place un tel service pour les raisons suivantes :

- contraintes de coûts (infrastructures matérielles et logicielles) ;
- environnement technique complexe (multi-sites) ;
- organisation multi-tutelle.

Certaines structures, comme l'Institut des Sciences Cognitives, qui offrent déjà le service à leurs personnels, mais dont les infrastructures arrivent en fin de vie prochainement, accueillent avec enthousiasme le projet porté par la DSI.

Compte tenu des impacts sur le réseau informatique d'un tel service, le CISR est également représenté au sein du groupe de travail.

Le groupe de travail se réunit une fois par mois, et les échanges sont possibles au travers d'une liste de diffusion et d'un espace collaboratif documentaire.

Spécifications du service proposé.

Public visé	Les personnels de l'université et les personnels hébergés de nos unités de recherche.
Risques supprimés ou diminués	• perte définitive de données • vol de données • cession de droits et de propriété intellectuelle sur des données professionnelles
Dimensionnement de la solution	• 1000 utilisateurs • 50Go en moyenne par utilisateur • Rétention des données maximale au regard des capacités de stockage (1 an semble un minimum - 5 ans demandés par certaines structures)
Performances et fonctions exigées	• la protection des données ne devra plus être à la charge de l'utilisateur mais administrée centralement par la DSI et les ASR/Correspondants locaux. • la protection devra être continue quelle que soit la situation géographique de l'utilisateur et gérer les différentes versions de fichiers. • la solution devra avoir une empreinte système minimale sur le poste ne perturbant pas le travail de l'utilisateur. • tous les postes de l'utilisateur (fixe, portable), quelle que soit la plate-forme (Windows, Mac OS, Linux), devront pouvoir être protégés
Ergonomie	• une mise en place sans configuration de la part de l'utilisateur. • un point d'accès unique fournira une vue consolidée de toutes les données sauvegardées de l'utilisateur quel que soit le poste informatique d'origine. • l'utilisateur pourra être autonome dans les tâches de recherche et de restauration de ses données.
Nouveaux usages	• accès et sauvegarde des données depuis des terminaux mobiles (iOS, Android) • en cas de vol ou de perte d'un terminal mobile, l'utilisateur pourra le géo-localiser et décider de supprimer les données à distance.
Exigences techniques	• la solution devra s'intégrer au système d'information de l'université, permettant à l'utilisateur de s'authentifier avec son identifiant et son mot de passe habituels. • les communications et le transfert des données sauvegardées se feront au travers d'un canal chiffré et sécurisé de bout en bout. • les données sauvegardées seront stockées au sein d'une infrastructure sécurisée, fiable et pérenne de la DSI • la gestion des stratégies de sauvegarde sera faite en collaboration avec les ASR/Correspondants offrant le service au sein des unités de recherche, afin de répondre aux besoins spécifiques de chacune.

Solutions à l'étude.

Le tableau ci-dessous énumère les solutions en cours d'étude par l'équipe Pédagogie et Recherche du pôle Infrastructures de la DSI et plusieurs informaticiens de laboratoire.

Nom de la solution	**Editeur**	**Orientation**	**Mode SaaS[2]** ***On demand***	**Hébergement sur site[3]** ***On premise***
Crash Plan	Code 42	Utilisateur	Oui	Oui
inSync	Druva	Utilisateur	Oui	Oui
Simpana	Commvault	Utilisateur	Oui	Oui
Live Navigator	ASG	Poste	Via un revendeur	Oui
Avamar	EMC	Poste	Via un revendeur	Oui

Les offres orientées *utilisateur* sont plus intéressantes que les offres orientées *poste*. Elles sont plus simples à utiliser pour un enseignant-chercheur ou un chercheur. Elles peuvent être utilisées de n'importe où (si on dispose d'une connexion Internet) et depuis n'importe quel poste de travail de l'utilisateur.

Leur disponibilité en mode SaaS impose une interface conviviale et simple pour l'utilisateur qui n'a pas de travail de configuration du logiciel à effectuer. Ces deux atouts du logiciel ont pour objectif que le client sollicite le moins possible le service d'assistance du fournisseur.

Les performances sont généralement bonnes, l'outil devant se plier aux contraintes de mobilité et de connectivité des utilisateurs. Il est assez simple de faire évoluer la solution pour supporter un plus grand nombre d'utilisateurs ou un volume de données sauvegardées plus important. Les utilisateurs peuvent s'authentifier avec leur identifiant et leur mot de passe habituels. La gestion d'une partie des données à sauvegarder peut être déléguée, par exemple à un informaticien en poste dans une unité de recherche.

Les offres orientées *poste* sont des solutions de sauvegarde plus traditionnelles dotées de fonctionnalité *poste de travail*. Leur fonctionnement, centré sur le poste et non sur l'utilisateur, apporte une couverture fonctionnelle moindre que leurs concurrents et des manipulations plus complexes pour l'utilisateur et les administrateurs dans des scénarios

[2] Software As A Service : Le client loue un service logiciel. Le logiciel est implémenté sur un serveur du fournisseur situé à l'extérieur de l'université. Les données sont aussi sauvegardées sur un site extérieur à l'université. On appelle aussi ce mode «*On demand*».

[3] Le client achète le droit d'usage non exclusif du logiciel et un contrat de maintenance logicielle. Le logiciel est implémenté sur un serveur de l'université. Les données restent à l'université. On appelle aussi ce mode «*On premise*».

particuliers (restauration croisée entre 2 postes, migration des données entre deux postes).

Le tableau ci-dessous donne une estimation des coûts sur la base de prix publics ou d'une première offre en prix remisé.

Nom de la solution	**Editeur**	**Coûts estimatifs**
Crash Plan	Code 42	• Prix public « on premise » : 56 euros par utilisateur et par an • Prix remisé constaté (INRIA Rennes en 2013) : 21 euros / utilisateur / an pour un volume de 225 licences
inSync	Druva	• Prix public « on premise » : 44 euros par utilisateur et par an • Première proposition commerciale : 40 euros / utilisateur /an pour un volume de 800 licences
Simpana	Commvault	• Prix public « on premise » : 46 euros par utilisateur et par an • Première proposition commerciale : 23 euros / utilisateur /an pour un volume de 1001 licences et engagement de 4 ans
Live Navigator	ASG	• Licence au poste, perpétuelle + maintenance annuelle • Prix via Groupe Logiciel : 30 euros par poste avec 3 ans de maintenance (pour 500 postes)
Avamar	EMC	Pas de proposition connue

Groupe de travail dédié au projet.

Le tableau ci-dessous donne la liste des informaticiens participant aux études des solutions du marché. Le chef de projet est Vincent HURTEVENT.

Amélie CORDIER	UMR5205 Laboratoire Informatique Images et Systèmes d'Information
Gilles BROCHET	UMR5205 Laboratoire Informatique Images et Systèmes d'Information
Yannick PERRET	UMR5205 Laboratoire Informatique Images et Systèmes d'Information
Laurent AZEMA	UMR5208 Institut Camille Jordan
Bruno SPATARO	UMR5558 Biométrie et biologie évolutive
Christophe PERA	Département Mécanique de la FST
Claude BONURA	UMR5246 Institut de Chimie et Biochimie Moléculaires et Supramoléculaires
Nicolas FOULON	UMR5246 Institut de Chimie et Biochimie Moléculaires et Supramoléculaires
Danis ABROUK	UMR5557 Ecologie Microbienne
Francisco PINTO	UMR 5306 Institut Lumière Matière
Said MEZZOUR	UMR 5306 Institut Lumière Matière
Hervé HUGUENEY	UMR_S1028 Centre de Recherche en Neurosciences de Lyon
Marie-Hélène LASSALLE	UMR5574 Centre de recherche astrophysique de LYON
Philippe FORTIN	UMR5276 Laboratoire de Géologie de Lyon : Terre, Planètes et Environnement
Sylvain MAURIN	Institut des sciences cognitives
Yoann LAFON	UMR_T9406 Laboratoire de biomécanique et mécanique des chocs
Yvan DOYEUX	UMR5270 Institut des Nanotechnologies de Lyon
Rémi SADER	CISR (Le réseau)
Vincent HURTEVENT	**DSI – pôle Infrastructures – Equipe pédagogie & recherche**
Jean-Michel EJENGELE	DSI – pôle Infrastructures – Equipe pédagogie & recherche
Adja Fatou FALL	DSI – pôle Infrastructures – Equipe pédagogie & recherche
Anne-Lyse PAPINI	DSI – pôle Infrastructures