



L'archivage des messageries électroniques

Preuve de concept VITAM

Version 1.2. – 31 octobre 2013

Équipe-projet VITAM :

Frédéric Brégier (MCC/SDSI)
Frédéric Deguilhen (MAE/DSI)
Lourdes Fuentes-Hashimoto (MAE/AR)
Marie Laperdrix (MCC/AN)
Thomas Van de Walle (MCC/AN)
Nathalie Morin (MINDEF/DMPA)
Edouard Vasseur (MINDEF/DMPA)

SUIVI DES MODIFICATIONS

Version	Action	Auteur	Date
v. 0.1.	Initialisation	Equipe-projet VITAM	08/03/2013
v. 1.0.	Diffusion au Comité interministériel aux Archives de France (CIAF)	Equipe-projet VITAM	27/06/2013
v. 1.1.	Intégration des remarques du ministère des Affaires étrangères, du ministère de la Culture et de la Communication (Archives nationales et SIAF) et du ministère de la Défense (DMPA et SHD)	Equipe-projet VITAM	03/09/2013
v. 1.2.	Consolidation de la version 1.1 suite à l'appel à commentaires de septembre 2013	Equipe-projet VITAM	31/10/2013

Table des matières

Table des matières	3
Table des illustrations.....	5
Introduction.....	6
Contexte	6
Objectifs et déroulement	7
I - Les messageries : contexte d'utilisation.....	8
1.1. Un outil technique	8
1.1.1. Définition et historique.....	8
1.1.2. Architecture globale.....	11
1.1.3. Le message électronique.....	12
Un objet encodé selon des standards spécifiques	12
Un objet composite	13
1.2. ...au croisement de plusieurs droits	14
1.2.1. Les différents droits applicables	14
Droit du travail.....	15
Droit de la preuve	15
Droit des archives et des documents administratifs.....	17
Droit au respect de la vie privée	17
1.2.2. Comment concilier les différentes contraintes juridiques ?.....	19
Transparence et proportionnalité au cœur des processus de contrôle.....	19
Règlement intérieur ou charte d'utilisation des équipements informatiques.....	20
Le rôle-clé des instances de consultation du personnel	20
Quelle approche pour le processus d'archivage ?.....	21
1.3. Une utilisation dépendante de l'organisation et de l'utilisateur	22
1.3.1. Un outil largement utilisé mais non sans conséquence sur l'organisation du travail.....	22
Une appropriation rapide et durable	23
Un impact sur les rythmes et le temps de travail	23
Une tendance à la surcharge informationnelle.....	24
Des frontières brouillées entre vie personnelle et vie professionnelle, entre domicile et lieu de travail	25
Les cadres, premiers touchés	26
1.3.2. Les facteurs qui conditionnent l'utilisation de la messagerie.....	27
Facteurs techniques.....	27
Facteurs fonctionnels	29
Facteurs comportementaux	30
Conclusion : une grande complexité pour l'archivage	31
II - Quelle approche archivistique de l'objet messagerie ?.....	33
2.1. Les réflexions engagées au niveau national et international.....	34
2.1.1. L'approche « pédagogique ».....	34
2.1.2. L'approche « collecte et conservation ».....	36
2.1.3. L'approche « diplomatique »	38
2.2. Bilan des expériences	38
2.2.1. Que collecter ?	39
Quelles messageries ?	39

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

Quels messages ?	40
Quelles informations ?	42
2.2.2. Comment et quand collecter ?	44
La collecte globale	44
La collecte ciblée	45
2.2.3. Comment traiter et conserver ?	47
Traitement intellectuel	47
Traitement technique	49
2.2.4. Comment restituer ?	52
Conclusion	53
III - Les expérimentations menées dans le cadre du volet 2 du POC VITAM	54
3.1. Présentation générale des tests réalisés	54
3.1.1. Objectifs généraux	54
3.1.2. Choix et sélection des boîtes mails	55
3.1.3. Choix et installation des outils	55
3.2. Résultats	56
3.2.1. Les outils	56
Aid4Mail et EmailChemy	56
CERP Email Parser et Xena	57
Les outils des bibliothèques open source	58
3.2.2. Les traitements	60
Conversion d'un format vers un autre	60
Extractions	62
3.2.3. L'intégrité des exports	64
Objectifs et descriptifs des tests	64
Résultats	65
Le cas des messageries officielles	66
3.3. Conclusions	66
3.3.1. Bilan par outil	67
3.3.2. Bilan par format	68
Conclusion et propositions	71
Propositions générales	71
Propositions organisationnelles	71
Propositions juridiques	72
Propositions techniques	73
Propositions archivistiques	74
Propositions dans le cadre du programme VITAM	75
Collecte des messageries	75
Constitution des SIP	76
Annexe 1 : Bibliographie	78
Annexe 2 : Réglementation et jurisprudence	85
Annexe 3 : Études de cas	88
Ministère des Affaires étrangères	88
Ministère de la Culture et de la Communication	89
Ministère de la Défense	91
Annexe 4 : Les outils utilisés dans le cadre des tests	94
Annexe 5 : Résultat des tests d'extraction des métadonnées	98
Annexe 6 : Glossaire	99

Table des illustrations

Fig. 1 : évolution des modes d'accès aux messageries électroniques (Source : Site internet Campaign Monitor, http://www.campaignmonitor.com/resources/will-it-work/email-clients/).....	10
Fig. 2 : étapes de transmission d'un message électronique (Source : Wikipedia)	11
Fig. 3 : les différents aspects de l'archivage d'une messagerie.....	33
Fig. 4 : schéma d'identification des messages à conserver (Source : Conseil international des archives)	41
Fig. 5 : modélisation de l'AIP correspondant à un message électronique (Source : Projet Testbed) .	51
Fig. 6 : Schéma de l'AIP correspondant à un compte de messagerie (Source : CERP, Email preservation workflow, 2010, p. 2.)	52
Fig. 7 : projet de schéma de représentation d'un SIP Compte de messagerie dans VITAM.....	77
Fig. 9 : modélisation des possibilités d'utilisation des messageries par un agent du ministère des Affaires étrangères	89
Fig. 10 : modélisation des possibilités d'utilisation des messageries par un agent du ministère de la Culture et de la Communication	91
Fig. 11 : modélisation des possibilités d'utilisation des messageries par un agent du ministère de la Défense.....	93

Introduction

Contexte

Les messageries électroniques constituent l'une des technologies les plus utilisées et les plus diffusées au monde. Le Radicati Groupe estimait ainsi en 2011 le nombre de comptes de messagerie électronique à 3,1 milliards en 2011, avec un taux de croissance de l'ordre de 7 % par an. 75 % correspondent à des comptes personnels, 25 % à des comptes professionnels. En 2011, chaque messagerie électronique gérait en moyenne 105 messages électroniques par jour : 72 messages reçus dont 14 pourriels et 33 messages envoyés¹.

Ce mode de communication est devenu incontournable, même si son utilisation au cours du temps a évolué grâce aux nouvelles fonctionnalités offertes et aux modes d'appropriation par les différents utilisateurs². La masse d'informations échangées par le biais de cet outil est fantastique. Et c'est bien une **place centrale que les messageries occupent au sein des systèmes informationnels**. Leur gestion constitue un enjeu organisationnel, sécuritaire, juridique, financier, technique mais aussi mémoriel pour toutes les organisations et notamment les organisations publiques³. En considérant ces différentes approches, quel archivage des messageries peut-on envisager ?

Depuis plus de dix ans, la communauté archivistique française comme étrangère a mené de nombreuses expérimentations sur cet outil identifié très tôt comme étant au cœur de la prise de décision par les organisations. Pour les professionnels de la gestion de l'information, le contrôle de cette masse d'informations constitue un enjeu de taille sur lequel, ces dernières années, des progrès significatifs ont été réalisés⁴. Plusieurs documents de préconisations techniques et de gestion ont été diffusés en 2008-2009 en France : directive des Archives de l'État en Belgique reprise dans une note d'information de la direction des Archives de France⁵, fiche conseil de l'Association des archivistes français, livre blanc de la FedISA⁶, livre blanc du CR2PA⁷, fiche sur les courriels du Guide DISIC sur l'archivage électronique⁸. L'objectif du présent document est de permettre aux

1 The Radicati Group. *Email Statistics Report 2011-2015*, disponible en ligne: <http://www.radicati.com/wp/wp-content/uploads/2011/05/Email-Statistics-Report-2011-2015-Executive-Summary.PDF> (consulté le 11 mai 2013).

2 TRAN Sébastien, « Quand les TIC réussissent trop bien dans les organisations : le cas du courrier électronique chez les managers », *Management & Avenir*, 2010, 34, p. 200-215

3 BALLET Philippe, RIETSCH Jean-Marc. *Conserver les courriers électroniques ? Ou comment résoudre la problématique de l'archivage des e-mails*, Paris : FedISA, 2008, p. 10.

4 PROM Christopher J. *Preserving Email. DPC Technology Watch Report 11-01 December 2011*, Londres : Digital Preservation Coalition, 2011, p. 1.

5 SOYEZ Sébastien, *Directives pour la gestion et l'archivage numérique des e-mails*, Archives générales du Royaume – Algemeen Rijksarchief, version 1.1, octobre 2009.

6 BALLET P., RIETSCH J.-M.. *Op.cit.*

7 CLUB DES RESPONSABLES DE POLITIQUE ET PROJETS D'ARCHIVAGE [CR2PA]. *L'archivage des mails ou Les utilisateurs face aux mails qui engagent l'entreprise. Livre blanc proposé par le groupe de travail "Archivage des mails" du CR2PA*, Paris : CR2PA, 2009.

8 DIRECTION INTERMINISTERIELLE DES SYSTEMES D'INFORMATION ET DE COMMUNICATION (DISIC). *Archivage électronique. Guide de bonnes pratiques et annexes*, DISIC, 2012, t. 2, p. 19-20. Disponible en

archivistes français de disposer d'une synthèse actualisée en français de ces travaux.

L'équipe-projet VITAM (Valeurs Immatérielles Transmises aux Archives pour Mémoire), réfléchissant à la conception d'une solution d'archivage électronique adaptée aux besoins des ministères chargés de l'archivage à très long terme des documents d'intérêt historique, statistique ou scientifique constituant la mémoire de l'État et de ses opérateurs, ne pouvait que s'interroger sur la collecte, la conservation et la communication des documents et données produits et reçus dans les messageries. Elle l'a fait en s'inscrivant dans la suite de ces recherches, en tirant parti des expériences menées pour définir les modalités d'entrée de ce type d'archives dans la solution d'archivage électronique VITAM.

Objectifs et déroulement

En réalisant cette preuve de concept relative aux messageries électroniques, l'équipe-projet VITAM poursuivait trois objectifs principaux :

- définir une stratégie de collecte, de traitement et de conservation des messageries électroniques adaptée aux différents contextes des partenaires du projet VITAM ;
- identifier les outils et fonctionnalités techniques nécessaires pour un traitement technique de ces messageries électroniques, en amont de leur versement dans la solution VITAM ;
- définir un modèle de métadonnées pour les messageries conforme à l'état de la réflexion de l'équipe-projet VITAM en matière de structuration des métadonnées, et permettant de générer des fichiers utilisables dans le cadre des expériences techniques conduites pour la preuve de concept menée au même moment autour des technologies NoSQL.

Pour parvenir à ces objectifs, l'équipe-projet VITAM s'est appuyée sur une revue de la littérature et des expériences nationales comme internationales existantes. Elle a également défini et mis en œuvre une série de tests techniques réalisés conjointement par le ministère de la Culture et de la Communication (sous-direction des systèmes d'information) et par le ministère de la Défense (Service historique de la Défense).

La preuve de concept s'est déroulée selon la chronologie suivante :

- mars 2013 : définition des objectifs et recensement de la bibliographie sur le sujet ;
- avril 2013 : revue de littérature et réalisation de la première série de tests ;
- mai-juin 2013 : synthèse de la revue de littérature, réalisation de seconde série de tests et rédaction de la synthèse.

Les termes techniques utilisés dans la présente étude font l'objet d'un glossaire en annexe 6

I - Les messageries : contexte d'utilisation

Les messageries électroniques sont avant tout **un outil technique de communication**. Utilisées au domicile comme sur le lieu de travail, elles véhiculent indifféremment et sans cloisonnement informations de nature privée et informations de nature professionnelle. Afin de conserver les messages les plus pertinents et engageants, notamment dans le monde professionnel, il est nécessaire de comprendre le fonctionnement et l'appropriation de l'outil par les utilisateurs.

Qu'est qu'une messagerie ? Qu'est-ce qu'un message ? Quel droit les archivistes ont-ils d'intervenir sur la messagerie professionnelle d'un agent, dès lors qu'il l'utilise également pour véhiculer des informations d'ordre privé ? En quoi le paramétrage et l'usage par chaque individu et/ou par chaque organisation (choix du titulaire du compte de messagerie : individu ? Entité ? Fonction ?)

1.1. Un outil technique ...

1.1.1. Définition et historique

Une messagerie électronique est un « outil de communication asynchrone entre deux ou plusieurs personnes ou organisations basé sur des réseaux électroniques ouverts et/ou privés permettant d'échanger des messages en format texte et des documents en format électronique »⁹. La commission de terminologie et de néologie en donne la définition suivante : « service permettant aux utilisateurs habilités de saisir, envoyer ou consulter en différé des courriels »¹⁰.

Les informations échangées par le biais de cet outil sont communément appelées messages ou courriers électroniques, e-mails ou courriels. Ceux-ci ne constituent pas un type de document en tant que tel, mais une catégorie technique susceptible de véhiculer de nombreux types de documents (courrier, note, compte rendu de réunion, bon de commande, etc.).

C'est en 1965 que Tom Van Vleck et Noël Morris se sont envoyés ce qui peut être considéré comme le premier message électronique au Massachusetts Institute of Technology (MIT) en utilisant une commande utilisant le *Compatible Time-Sharing System* (CTSS)¹¹. La première messagerie électronique collective a été créée à la fin des années 1960 dans le cadre du projet NLS (oN-Line System), étroitement lié à la naissance du premier réseau distribué ARPANET en 1969.

En mars 1972, un ingénieur de la société BBN, Ray Tomlinson, écrit le premier programme de courrier électronique entre deux machines, à partir de deux programmes qu'il vient de créer pour le courrier « intra-machine » : SNDMSG (pour SeND MeSsaGe), pour envoyer des messages et

9 TRAN Sébastien, « Quand les TIC réussissent trop bien dans les organisations : le cas du courrier électronique chez les managers », *Management & Avenir*, 2010, 34, p. 200-215.

10 « Vocabulaire du courrier électronique », *Journal officiel de la République française*, 20 juin 2003. Les termes de courrier électronique, email et courriel sont utilisés indifféremment dans la suite de ce document.

11 PROM C. J. *Op.cit.*, p. 2.

README, pour lire les messages. La création d'un protocole expérimental de transfert de fichiers (appelé CPYNET) permet aux messages électroniques de « sortir » pour la première fois des machines locales pour circuler sur le réseau, créant ainsi le courrier électronique à distance. Ray Tomlinson choisit le caractère @ pour séparer l'adresse de l'utilisateur de celle de l'hébergeur. Le courrier électronique devient très vite la première application d'ARPANET utilisée par les chercheurs universitaires, alors que le réseau avait été conçu pour le partage des ressources informatiques. Dès 1973, une étude menée par l'ARPA montre que 75 % du trafic de son réseau provient des échanges de messages électroniques¹².

En 1975, le premier client de messagerie, MSG, est développé par John Vittal, ingénieur à l'Institut des sciences de l'information de l'Université de Californie du Sud. Son programme, considéré comme l'ancêtre des clients de messagerie modernes comme Outlook ou Thunderbird, permet à lui seul de rassembler les fonctions de lecture, d'expédition, de transfert des messages, d'ajout de pièces jointes ainsi que la notion de « corbeille » pour les messages supprimés, le tout dans une interface simplifiée.

Le début des années 1980 marque une nouvelle étape avec l'apparition en 1982 du protocole SMTP (RFC 821), encore limité aux simples textes encodés en ASCII, avec l'adoption par ARPANET, en 1983, du protocole TCP-IP (Transfert Control Protocol – Internet Protocol) puis avec l'ouverture progressive d'Internet et le développement des ordinateurs personnels (premiers IBM PC/AT sur processeurs Intel X86 en 1982). De nouveaux clients de messagerie sont alors développés : Eudora (Université de l'Illinois en 1985 par Steve Dorner), Outlook (Microsoft en 1997). Les messageries connaissent alors une expansion sans précédent que l'arrivée des réseaux sociaux, au début des années 2000¹³ n'a pas réussi à remettre en cause.

Les solutions de messages électroniques ont évolué avec les technologies :

- ⤴ les premières versions des années 1970-1980 étaient limitées à des solutions utilisant la notion de client « passif », car les modèles d'ordinateurs étaient alors des terminaux passifs connectés à un serveur central (*Mainframe*). Les messages étaient alors stockés sur le serveur central ;
- ⤴ suite à la diffusion des ordinateurs personnels à la fin des années 1980, des solutions en clients « lourds » sont progressivement apparues dans les années 1990 (Outlook, Thunderbird, ...). Il était désormais possible de stocker ses messages sur le disque dur des ordinateurs individuels ;
- ⤴ l'explosion d'Internet à la fin des années 1990 et au début des années 2000 a favorisé la création de service de messageries en ligne (Gmail par Google en 2004), redonnant une certaine liberté de mouvement aux usagers pour lire leur messagerie *via* des clients « web ». Les messages peuvent à nouveau être stockés sur un serveur central, avec l'utilisation du protocole HTTP entre le client et le serveur, et le protocole SMTP entre les serveurs de messageries ;
- ⤴ l'apparition des ordiphones et autres solutions de mobilité (tablettes) à la fin des années 2000 a permis l'apparition d'un modèle intermédiaire, le client « mobile », entre le modèle « passif » des années 1980 et le modèle « web ». Le stockage des messages n'est plus sur le

12 <http://www.sites.univ-rennes2.fr/urfist/ressources/utilisation-avancee-de-la-messagerie-electronique/principes-de-fonctionnement-de-la-messa> et http://www.demainlemail.com/email_alinto/historique/ , <http://www.irepp.com/blog/2006/03/15/un-breve-histoire-de-le-mail/> (consultés le 11 mai 2013).

13 http://www.demainlemail.com/email_alinto/historique/ , <http://www.irepp.com/blog/2006/03/15/un-breve-histoire-de-le-mail/>(consulté le 11 mai 2013).

client mais sur le serveur central (comme pour le modèle « web »), mais le modèle est définitivement « mobile », au sens où les contraintes de déplacement de l'utilisateur sont balayées. Le protocole utilisé est l'IMAP entre le client (mobile) et le serveur, les serveurs utilisant toujours le protocole SMTP.

Plusieurs études menées récemment ont montré une accélération du passage de l'usage de la messagerie des clients lourds (Outlook, Thunderbird) vers les clients mobiles ou web. Depuis le début de l'année 2012, les clients mobiles dépassent les clients lourds qui se retrouvent en nette perte de vitesse, tandis que les clients web se maintiennent¹⁴.

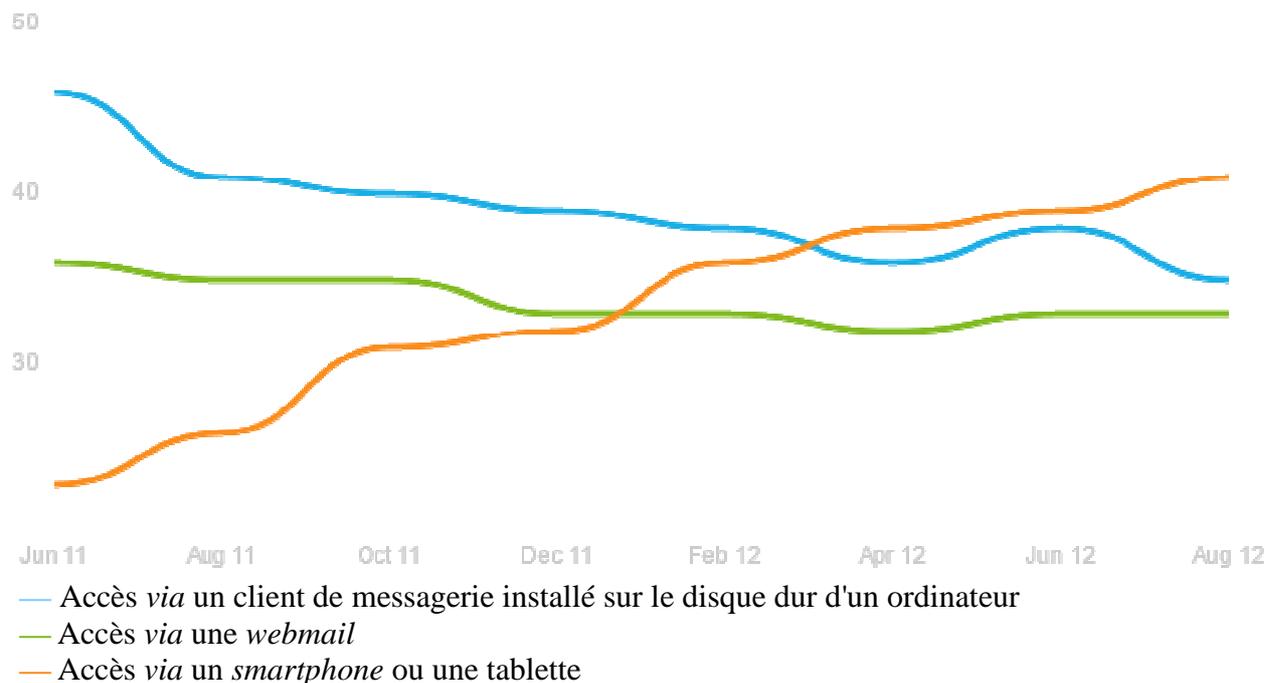


Fig. 1 : évolution des modes d'accès aux messageries électroniques (Source : Site internet Campaign Monitor, <http://www.campaignmonitor.com/resources/will-it-work/email-clients/>)

Cette tendance est confirmée par la répartition des outils ou méthodes d'accès aux messages dans les mêmes études. Parmi les clients lourds, le logiciel Outlook de Microsoft confirme sa position de *leader* du marché.

Devant la multiplicité des outils, des moyens d'accès, et même des types de messageries, certaines entreprises ont décidé de supprimer les messages électroniques en les remplaçant par des réseaux sociaux (Atos Origin en 2011). Les protocoles sont alors principalement http. Le protocole SMTP disparaît. D'autres ont décidé d'implanter un système « mobile » et « ordonné », sur la base du principe conjoint d'un réseau social et d'un client de messagerie. Il s'agit par exemple du projet Diplomatie mené par le ministère des Affaires étrangères. Le protocole est principalement http, tout en prenant en compte les clients mobiles.

14 Sources : études Campaign Monitor de 2011 et 2012 (disponibles en ligne : <http://h30565.www3.hp.com/t5/Articles-de-nos-correspondants/Webmails-ou-clients-de-messagerie-quelle-est-la-meilleure/ba-p/3513?profile.language=fr>), étude IDC "Always Connected" de mars 2013 (disponible en ligne : <https://fb-public.box.com/s/3iq5x6uwnqtq7ki4q8wk>).

1.1.2. Architecture globale

La messagerie électronique s'appuie principalement sur le réseau Internet pour assurer le transfert d'information selon des standards spécifiques de routage et de contenu entre deux utilisateurs. Le **support de ce transfert est appelé courrier électronique** ou e-mail ou courriel.

Ce système de transfert d'information demande que les utilisateurs disposent :

- ♣ d'une adresse électronique : l'adresse électronique est une chaîne de caractères constituée de l'identifiant de l'utilisateur et de celui du gestionnaire d'un service de messagerie électronique séparés par le caractère @, identifiant de manière unique cet utilisateur ;
- ♣ d'un accès à internet ;
- ♣ d'un client de messagerie ou d'un *webmail*.

Le transfert s'effectue sur le réseau internet sur des serveurs spécifiques et selon des standards propres.

Autant la transmission des messages électroniques est normalisée autour de standards ouverts, autant le processus entourant la réception et le stockage des messages est faiblement normalisé¹⁵.

Le courrier électronique transite de serveurs en serveurs jusqu'au destinataire.

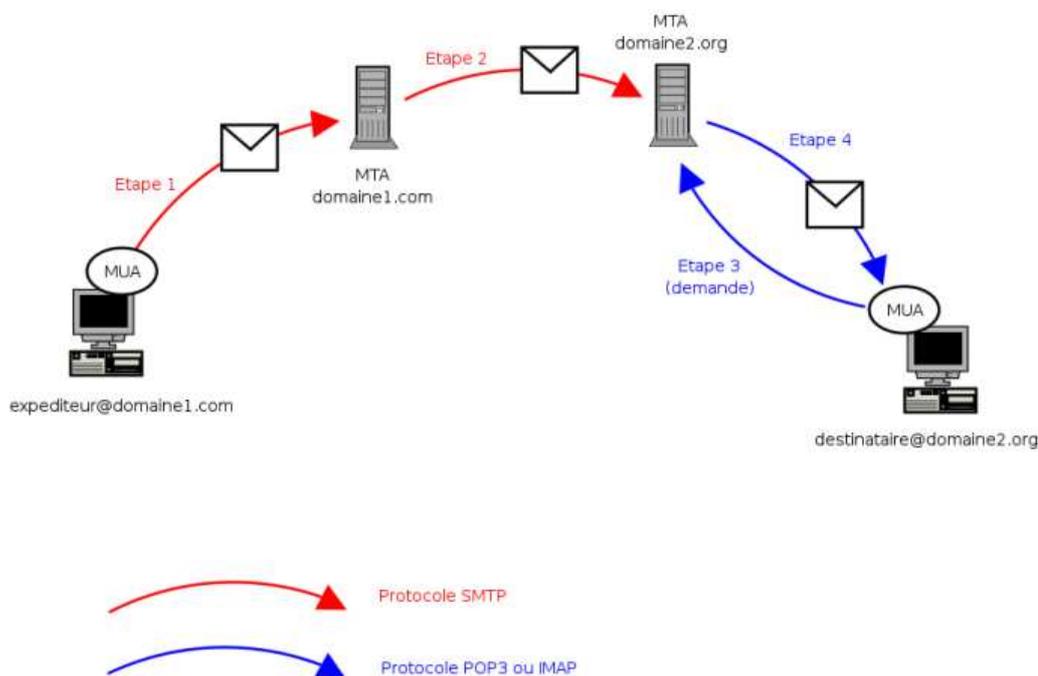


Fig. 2 : étapes de transmission d'un message électronique (Source : Wikipedia)

Le message est d'abord envoyé au serveur de transfert (Mail Transport Agent MTA) de l'expéditeur, *via* son réseau local ou *via* son réseau internet. Ce serveur achemine le message au

15 PROM C. J. *Op.cit.*, p. 19-20.

serveur MTA du destinataire *via* le réseau internet. Ces serveurs communiquent entre eux à l'aide du protocole SMTP (Simple Message Transfert Protocol). Le serveur MTA du destinataire envoie alors le message au serveur de messagerie entrant ou MDA (Mail Delivery Agent) qui stocke le courrier jusqu'à ce que le destinataire vienne le relever.

Il existe plusieurs protocoles permettant de relever le courrier dont le fonctionnement détermine la localisation de stockage des messages :

- ♣ le protocole POP3 (Post Office Protocol) : les messages sont relevés du serveur et directement stockés sur l'équipement de communication (*device*) de l'utilisateur ;
- ♣ le protocole IMAP (Internet Message Access Protocol) : il permet une synchronisation de l'état des courriers entre plusieurs clients de messagerie. Les messages restent stockés sur le serveur tout en pouvant être dupliqués sur l'équipement local¹⁶.

L'envoi et la relève du courrier électronique se font à l'aide d'un logiciel spécifique appelé MUA (Mail User Agent). Lorsque celui-ci est installé sur le poste de l'utilisateur, on parle de client de messagerie (Outlook, Thunderbird, Lotus note par exemple). Lorsque l'accès à la messagerie se fait *via* une interface web, on parle de *webmail* (hotmail, yahoo mail, gmail, *webmails* des principaux fournisseurs d'accès internet)¹⁷.

On peut considérer que les serveurs MTA font office de bureau de poste où sont effectuées les tâches de prise en charge du courrier, de tri et d'acheminement, le protocole SMTP joue le rôle de facteur, le serveur MDA de boîte aux lettres pour stocker les messages.

1.1.3. Le message électronique

Un objet encodé selon des standards spécifiques

Pendant les quinze premières années de leur utilisation, les messages électroniques n'ont pas fait l'objet d'un format normalisé. Ce n'est qu'à partir de 1981 que le Network Working Group de l'Internet Engineering Task Force (IETF) a commencé à préciser les méthodes pour envoyer et recevoir des messages. Les caractéristiques techniques utilisées dans les messageries électroniques sont désormais définies dans une série de « *request for comment* » (RFC), littéralement « demandes de commentaires », publiées par l'IETF.

Historiquement les messages électroniques ne contenaient que du texte encodé en ASCII 7-bits (alphabet sans accent), seul encodage supporté par le protocole SMTP. De plus, le protocole SMTP imposait certaines restrictions sur le contenu des messages : composés uniquement de caractères ASCII 7-bits, les lignes ne pouvaient pas excéder 1000 caractères et la taille totale du contenu ne pouvait pas excéder une certaine dimension.

Avec le développement du multimédia et de l'usage des applications bureautiques est apparu le besoin d'envoyer des fichiers binaires (son, image, fichiers compressés, fichiers bureautiques). Le standard MIME (Multipurpose Internet Mail Extension) étend le format de données des courriels pour leur permettre de supporter des textes en différents codages de caractères et des contenus non

¹⁶ *Ibid.*, p. 21.

¹⁷ SOYEZ S., *Op. cit.*, p. 20.

textuels. MIME prend en charge le message électronique. Il en encode les différentes parties (texte principal, pièces jointes) et place dans l'en-tête les informations pour que le logiciel qui réceptionne le message puisse le décoder. Il convient de noter que certains modes de codage (notamment le codage en base 64) induisent une augmentation conséquente du volume des données envoyées (jusqu'à + 30 %)¹⁸.

L'information est encodée au départ et décodée à l'arrivée.

La **RFC 2822 spécifie le format des en-têtes et du corps du message**, ainsi que les règles d'en-têtes générales comme des champs To :, From : et Date :. Le standard MIME définit un ensemble d'attributs additionnels aux en-têtes pour le type de contenu et le codage, ainsi que les règles spécifiques pour encoder les caractères non ASCII dans les en-têtes. Le contenu du corps du message était initialement réglementé par la RFC 2821¹⁹.

Chaque logiciel de messagerie offre la possibilité de sauvegarder les messages dans un format d'export (soit propriétaire, soit ouvert). La plupart des logiciels en utilisent plusieurs (cf. Outlook qui permet des exports aux formats .pst, .msg, .html), tandis que certains ne sont utilisés que par un seul logiciel (cf. Lotus et format .nsf). Ces formats d'export permettent ou non de conserver l'intégralité des informations tant de l'en-tête que du corps²⁰.

La normalisation de certaines parties des messages, et notamment de l'en-tête, ainsi que leur stockage sur des serveurs centralisés où ils ne peuvent être modifiés par les utilisateurs constituent des atouts pour l'archivage. L'utilisation croissante des serveurs pour le stockage et le recours à des services de messagerie appuyés sur un stockage en mode *cloud computing* facilitent également la capture des messages à partir d'un point centralisé²¹.

Un objet composite

Les messages électroniques permettent aux utilisateurs d'envoyer tout type d'information numérique, y compris des fichiers produits avec d'autres applications. Ils constituent des objets complexes, formés de multiples composants :

▲ des données structurées : l'en-tête ;

▲ une masse plus importante de données non structurées : le corps du message et les pièces jointes²².

L'en-tête comprend les informations relatives à la rédaction, à l'expédition, à la transmission et à la gestion des messages : identité de l'expéditeur ; identité des destinataires (destinataires principaux, destinataires pour information, destinataires cachés) ; objet ; dates (envoi, réception par les différents serveurs intermédiaires, réception) ; niveau de priorité et de sensibilité ; accusé de réception ; suivi (indexation relative à l'action à réaliser par le destinataire)²³.

18 *Ibid.*, p. 17.

19 *Ibid.*, p. 18.

20 *Ibid.*, p. 11.

21 PROM C. *Op.cit.*, p. 12.

22 *Ibid.*, p. 8-9.

23 DIGITAL PRESERVATION TESTBED. *From digital volatility to digital permanence: Preserving email*. The Hague, avril 2003, p. 22-23 et SOYEZ S., *Op. cit.*, p. 16

Le corps du message constitue son contenu à proprement parler. Il est plus ou moins rédigé, structuré et mis en forme, selon la volonté du rédacteur et son choix de mise en page. Il peut intégrer des liens hypertextes vers des ressources extérieures (fichiers présents sur le réseau ou ressources disponibles sur Internet) ou des fichiers multimédia (photographies, fichiers vidéo)²⁴.

Les pièces jointes sont des fichiers dans n'importe quel format, avec pour seule limite la taille maximale autorisée pour l'envoi de fichiers²⁵.

1.2. ...au croisement de plusieurs droits²⁶

Plusieurs types de réglementations influent sur la gestion et l'archivage des messageries électroniques, notamment dans le cadre de leur utilisation dans un contexte professionnel. Les principales sont celles ayant trait au droit du travail, au droit de la preuve, au droit du patrimoine et au droit au respect de la vie privée. Prises ensemble, ces réglementations ont encouragé une attitude passive face à l'archivage des messageries électroniques²⁷.

1.2.1. Les différents droits applicables

Les messages électroniques n'ont reçu une définition juridique qu'en 2002, avec la publication de la directive européenne « vie privée et communications électroniques » du 12 juillet 2002 : « tout message sous forme de texte, de voix, de son ou d'image envoyé par un réseau public de communication qui peut être stocké dans le réseau ou dans l'équipement terminal du destinataire jusqu'à ce que ce dernier le récupère ». Cette définition couvre également les services de messagerie courts (SMS) et les services de messagerie multimédia (MMS)²⁸.

Il a fallu deux ans pour que la directive soit transposée en droit français par la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique (LCEN) qui, dans son article 1, précise : « on entend par courrier électronique tout message, sous forme de texte, de voix, de son ou d'image, envoyé par un réseau public de communication, stocké sur un serveur du réseau ou dans l'équipement terminal du destinataire, jusqu'à ce que ce dernier le récupère »²⁹.

Plusieurs réglementations s'affrontent autour de la question des messages électroniques, faisant d'eux des objets juridiques d'une grande complexité. Le principal problème réside naturellement dans la conciliation des intérêts organisationnels protégés par le code du travail et le code du patrimoine (continuité de service, archivage) et du droit au respect de la vie privée.

24 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 24 et SOYEZ S., *Op. cit.*, p. 17

25 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 25.

26 ARCHIVES NATIONALES, *La gestion et l'archivage des courriels, manuel pratique*, version 2, août 2012, p. 47-50.

27 PROM C. J., *Op. cit.*, p. 13-14.

28 PREVOT Maxime, *L'archivage des courriels, constat juridique d'une procédure négligée*. Mémoire de l'université de Paris Sud XI sous la direction de Marie Cornu et Hélène Guichard-Spica, 2012, p. 15.

29 BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 11 et PREVOT M., *Op. cit.*, p. 15.

Droit du travail

Le juge a reconnu une **présomption du caractère professionnel des dossiers et fichiers créés par un salarié grâce à l'outil informatique mis à sa disposition par son employeur**, sauf identification comme personnels par l'employé³⁰. Ce principe s'applique également aux messages électroniques, de sorte que l'employeur est en droit de les ouvrir hors la présence de l'intéressé, sauf s'ils sont identifiés comme personnels³¹. Le juge administratif a présumé de manière identique pour les messages électroniques rédigés par les fonctionnaires sur leur messagerie professionnelle, sauf à ce que leur contenu intéresse de manière évidente la vie privée de son auteur³².

En raison de cette présomption de nature professionnelle pour les messages et dossiers électroniques, l'employé est tenu, lorsque l'employeur en fait la demande, de restituer les éléments matériels et de communiquer les informations qu'il détient et qui sont nécessaires à la poursuite de l'activité de son entreprise³³. Chiffrer volontairement l'accès à ses données sur son poste informatique, sans l'autorisation de son employeur, constitue pour un employé une faute grave³⁴.

Cette présomption de la nature professionnelle des fichiers et messages conduit logiquement à retenir que les **fichiers personnels doivent être expressément identifiés comme tels par l'employé**. La définition du caractère personnel ne peut cependant être trop large³⁵. Ainsi, le fait de nommer un fichier d'après ses initiales ne peut lui conférer de caractère personnel³⁶.

Droit de la preuve

En droit français, la preuve appartient à un régime dualiste. Le recours implique, selon les circonstances, la preuve par écrit ou la preuve libre.

La preuve par écrit, pour tous les litiges mettant en cause une somme supérieure ou égale à 1500 euros, relève des dispositions du code civil (art. 1341). Un écrit non signé peut avoir un effet probatoire, soit en tant que commencement de preuve par écrit (code civil, art. 1347-1348), soit en tant que copie (code civil, art. 1348), soit en cas d'impossibilité de prouver. La copie a une force juridique variable selon les cas. Reproduction littérale d'un original, sans être revêtue des signatures qui en feraient un second original, elle ne fait foi que lorsque celui-ci n'existe plus (code civil, 1348). La copie doit être une reproduction fidèle et durable (code civil, art. 1348). En cas de contestation d'un écrit ou d'une signature électronique, l'article 287 du code de procédure civile dispose que le juge doit vérifier si les exigences des articles 1316-1 et 1316-4 du code civil sont satisfaites. L'article 1316-1 du code civil dispose que l'écrit électronique est recevable à titre de preuve, au même titre que l'écrit papier, sous réserve :

- qu'au moment de son établissement, la personne dont il émane puisse être dûment identifiée et qu'il soit établi dans des conditions de nature à en garantir l'intégrité ;

30 Soc. 18 octobre 2006 et BALLETT P., RIETSCH J.-M., *Op. cit.*, p. 13.

31 Soc. 6 juin 2012 et PREVOT M., *Op. cit.*, p. 30-31.

32 CAA Rennes, 14 janvier 2010.

33 Soc. 6 février 2001.

34 Soc. 18 octobre 2006.

35 Soc. 10 juin 2008.

36 Soc. 21 octobre 2009.

- qu'il soit ensuite conservé dans des conditions de nature à en garantir l'intégrité³⁷.

Depuis l'ordonnance n° 2005-674 du 16 juin 2005 relative à l'accomplissement de certaines formalités contractuelles par voie électronique, l'utilisation de la voie électronique est par ailleurs permise pour mettre à disposition des conditions contractuelles ou des informations sur des biens et des services (code civil, art. 1369-1) et permet de transmettre par courrier électronique des informations pour la conclusion d'un contrat si le destinataire l'a accepté (code civil, art. 1369-2). Le professionnel chargé du contrat peut désormais être contacté par voie électronique (code civil, art. 1369-3).

La preuve libre, de son côté, s'obtient par tous les moyens, notamment en droit administratif. L'écrit n'est pas nécessaire pour obtenir la force probante même s'il est prépondérant en pratique³⁸. L'administration fiscale ne considère pas un courrier électronique émis par un de ses agents comme un document portant interprétation d'un texte fiscal et, à ce titre, opposable en justice³⁹.

S'agissant des messages électroniques, le juge judiciaire est singulièrement mal à l'aise et dénie fréquemment toute valeur de preuve aux messages électroniques, pour de multiples raisons : doute sur l'auteur des messages⁴⁰ ; collecte des messages dans des circonstances impropres à en garantir l'authenticité⁴¹ ; non satisfaction des conditions fixées par l'article 1316 du code civil⁴² ; possibilité de créer de toute pièce, de modifier ou d'antidater un message⁴³. Au mieux, il considère que les messages électroniques constituent des commencements de preuves⁴⁴.

Quant au juge administratif, il a refusé de reconnaître la valeur probante d'un message

37 BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 14.

38 PREVOT M., *Op. cit.*, p. 9-10.

39 SJ - Garantie contre les changements de position de l'administration fiscale - Garantie contre les changements de doctrine - Procédures de rescrit fiscal - Garantie contre les changements d'interprétation d'un texte fiscal - Garantie apportée par une prise de position formelle au profit d'un contribuable sur l'interprétation d'un texte fiscal, consultable à l'adresse : <http://bofip.impots.gouv.fr/bofip/616-PGP.html?identifiant=BOI-SJ-RES-10-10-20-20120912> (consulté le 11 mai 2013).

40 La Cour d'appel de Versailles, dans un arrêt du 29 janvier 2004, a dénié toute force probante à des messages produits à partir de *webmails*, étant donné qu'aucun justificatif d'identité n'est demandé lors de la création du compte. CA Versailles, 2^e ch., 29 janvier 2004, pourvoi n°03-01521. Ce doute sur l'imputabilité des envois a été confirmé par un arrêt de la Cour d'appel de Paris en date du 9 novembre 2005 (CA Paris, 22^e chambre, 9 novembre 2005, pourvoi n° 04-36631).

41 Le 24 juin 2009, la chambre sociale de la Cour de cassation a dénié la valeur probante d'un message électronique qu'elle considérait recueilli dans des circonstances impropres à en garantir l'authenticité et selon un cheminement informatique non clairement défini, sans identification claire de l'auteur. Soc. 24 juin 2009. PREVOT M., *Op. cit.*, p. 17-18.

42 Le 30 septembre 2010, la même chambre a cassé un arrêt aux motifs que, dans une affaire de résiliation de bail, la cour d'appel n'avait pas recherché si les conditions mises par l'article 1316 du code civil étaient satisfaites. Le juge aurait dû constater que les courriers électroniques concernés n'étaient pas revêtus d'une signature électronique répondant aux conditions fixées par le code civil. Le courrier électronique simple, non revêtu d'une signature électronique, ne constitue pas un écrit sous forme électronique et ne peut être admis comme preuve. 1^e civ., 30 septembre 2010. PAPIN E., *Op. cit.*, p. 38-39.

43 Le 22 mars 2011, la Cour de cassation a estimé qu'un message électronique pouvait très bien être modifié ou créé de toute pièce et antidaté, lui déniant ainsi toute valeur probante. Soc. 22 mars 2011, pourvoi n° 09-43307. PREVOT M., *Op. cit.*, p. 18.

44 Le 20 mai 2010, la première chambre civile, dans un litige lié à l'existence d'un bon de commande, a reconnu que les messages électroniques échangés entre parties ne constituaient que des commencements de preuve par écrit. 1^e civ., 20 mai 2010. PREVOT Maxime M., *Op. cit.*, p. 18 et PAPIN Étienne, « La jurisprudence sur le courrier électronique », *Archimag*, décembre 2010-janvier 2011, n° 240, p. 38-39.

électronique stocké dans la partie « message envoyé » d'un client de messagerie⁴⁵.

Droit des archives et des documents administratifs

Les archives sont définies par le code du patrimoine, reprenant la loi n° 79-18 du 3 janvier 1979 modifiée sur les archives, comme « l'ensemble des documents, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité » (art. L211-1). Si ces documents ont été produits ou reçus dans le cadre d'une mission de service public, ils constituent des archives publiques au sens de l'art. L211-4 du même code. S'attachent en conséquence à eux les obligations qui pèsent sur les archives publiques en application du code du patrimoine : imprescriptibilité (art. L212-1), inaliénabilité (art. L212-1), contrôle scientifique et technique de l'Etat (destruction, externalisation, art. L212-2 à L212-5), libre communicabilité sauf présence d'informations dont la communication est protégée pendant une certaine durée par la loi (code du patrimoine, art. L213-1 et L. 213-2 et loi n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public dite loi CADA).

Les messages électroniques, dès lors qu'ils sont produits ou reçus par une personne dans l'exercice de son activité, **constituent de facto des archives au sens de la loi. S'ils ont été produits ou reçus par une personne chargée de tout ou partie de l'accomplissement d'une mission de service public, ils prennent la nature d'archives publiques et sont soumis aux mêmes obligations où qu'ils se trouvent.** C'est le lien entre le document et l'activité exercée qui est essentiel au regard de la réglementation.

La présence, au sein d'une même messagerie, de documents relevant d'un côté d'activités personnelles et de l'autre d'activités professionnelles, qu'elles relèvent d'une mission de service public ou non, rend la question de l'archivage des documents professionnels complexe⁴⁶. Celui-ci doit en effet être réalisé en garantissant la protection de la vie privée des salariés. En conséquence, toute solution d'archivage se fondant sur la collecte globale et automatique de l'ensemble des messages reçus par un salarié risque fort en effet d'être rendue impossible du fait du non respect des principes de transparence et de proportionnalité (voir ci-dessous). De même, un salarié risque de ne pas intégrer dans le système d'archivage un message mêlant vie professionnelle et vie personnelle, par crainte que la communication de ce document ne lui porte atteinte⁴⁷.

Droit au respect de la vie privée

Le **respect de la vie privée** est en effet au centre de la problématique de l'archivage des messages électroniques, en raison de l'utilisation conjointe du même outil pour échanger des correspondances d'ordre personnel et d'ordre professionnel.

L'article 8 de la Convention européenne des droits de l'homme de 1955 rappelle que « toute

45 *Ibid.*, p. 11.

46 Cette question se pose essentiellement pour les messageries interpersonnelles. Les messageries d'autorité et les messageries fonctionnelles sont censées ne contenir que des messages professionnels.

47 PREVOT M., *Op. cit.*, p. 49-50.

personne a droit au respect de sa vie privée et familiale, de son domicile et de sa correspondance ». Cette disposition s'impose à tous, et notamment à l'employeur, qui est tenu de respecter cette disposition sur le lieu de travail, comme l'a confirmé la Cour européenne des droits de l'homme (arrêt Niemietz, 16 décembre 1992). Le Code civil français, depuis la loi du 17 juillet 1970, reconnaît dans son article 9 à chacun un droit au respect de sa vie privée⁴⁸.

Le droit au respect de la vie privée peut cependant faire l'objet de limitations qui ne sont cependant acceptables qu'aux conditions de :

- ▲ **légalité** : ingérence prévue par une norme. La norme doit être claire, précise et accessible. Un règlement ou un contrat de travail peuvent suffire. Le consentement pourra dans certains cas être requis ;
- ▲ **finalité** : but légitime comme la prévention des infractions, le respect de la morale ou des droits et libertés d'autrui. Une charte visant à prévenir la fraude et la corruption d'un système informatique peut être considérée comme légitime ;
- ▲ **proportionnalité** : l'ingérence doit être limitée au strict nécessaire. La mesure doit être proportionnée au but poursuivi. Il faut vérifier si le but poursuivi par l'ingérence ne peut pas être atteint par des mesures non intrusives⁴⁹.

La jurisprudence a précisé que ce respect de la vie privée sur le lieu de travail comprenait le **respect du secret des correspondances et leur inviolabilité**⁵⁰. Des messages échangés entre collègues peuvent se voir reconnaître un caractère privé et être en conséquence couverts par ce principe de secret de la correspondance, notamment en cas de dénigrement d'un supérieur⁵¹.

L'employeur ne peut ainsi pas prendre connaissance des messages personnels émis par le salarié et reçus par lui grâce à un outil informatique mis à sa disposition pour son travail, sauf à s'exposer aux sanctions prévues par l'article 226-15 du Code pénal, même en cas d'interdiction d'un usage de la messagerie à des fins non professionnelles⁵². L'employeur ne peut prendre connaissance des fichiers identifiés comme personnels par le salarié et contenus sur le disque dur de l'ordinateur mis à sa disposition qu'en présence de l'agent, sauf risque ou événement particulier⁵³. Le juge a notamment autorisé un huissier à accéder aux fichiers sur le poste d'une employée aux motifs que l'employeur avait des raisons légitimes et sérieuses de craindre que l'ordinateur mis à disposition avait été utilisé pour favoriser des actes de concurrence déloyale⁵⁴.

Mais la notion de messages et fichiers personnels reste une notion mal définie, le législateur n'ayant donné aucune définition claire de la notion de message électronique personnel, malgré une tentative dans un projet de loi en date du 13 juin 2006⁵⁵. Les juges estiment que le formalisme dans la messagerie professionnelle du salarié permet de faire la distinction entre vie privée et vie professionnelle et qu'il y a une présomption de caractère professionnel à défaut d'identification par

48 PREVOT M., *Op. cit.*, p. 26-27.

49 ROBERT Romain, « Archivage des e-mails et protection de la vie privée », dans DEMOULIN Marie [dir.], *L'archivage électronique et le droit*, Bruxelles : Larcier, 2012. 195 p. (collection du CRIDS), p. 55-103.

50 Soc. 2 octobre 2001 (arrêt Nikon).

51 Soc. 6 juin 2007.

52 BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 13 et FERALE-SCHUHL Christiane. *Cyberdroit. Le droit à l'épreuve de l'internet*. Paris : Dalloz, 2010. 6^e édition.

53 Soc. 17 mai 2005 et BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 13

54 Soc. 10 juin 2008 et FERALE-SCHUHL C., *Op. cit.*

55 PREVOT M., *Op. cit.*, p. 16 et 29 et BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 11.

le salarié. En cas de doute, le juge peut être amené à scruter le contenu du message⁵⁶.

Le principe de l'inviolabilité s'applique également aux employés qui procèdent à une surveillance de leurs anciens supérieurs, après leur départ de l'entreprise. Le simple fait de consulter les courriers électroniques de tiers en utilisant ses codes d'accès constitue un accès frauduleux à un système informatique et une atteinte au secret des correspondances. La Cour de cassation a condamné une tentative de connexion au poste informatique du directeur d'une société⁵⁷.

1.2.2. Comment concilier les différentes contraintes juridiques ?

Concilier les différentes contraintes juridiques, et notamment l'articulation entre respect de la vie privée et pouvoir de contrôle des employeurs, a constitué jusque-là la priorité des pouvoirs publics. Sur la base de la réglementation et de la jurisprudence, la Commission nationale de l'informatique et des libertés (CNIL) a ainsi diffusé un certain nombre de recommandations, principalement destinées aux acteurs du secteur privé⁵⁸.

*Transparence et proportionnalité au cœur des processus de contrôle*⁵⁹

Les principes de transparence et de proportionnalité doivent être au cœur des actions des employeurs. Si la CNIL estime que des systèmes de contrôle de la messagerie électronique peuvent être mis en place par les entreprises pour des exigences de sécurité, de prévention ou de contrôle de l'encombrement du réseau (taille et volume de messages échangés de manière globale ; format et taille des pièces jointes), elle insiste également sur le fait que de tels systèmes peuvent être considérés comme une atteinte à la vie privée. Ces systèmes doivent naturellement faire l'objet d'une déclaration auprès de la CNIL, notamment dès lors que l'employeur veut mettre en place un contrôle poste par poste⁶⁰. L'employeur doit donc rechercher un juste équilibre entre pouvoir de contrôle et protection des libertés fondamentales des employés.

Principe de transparence : la jurisprudence insiste sur le fait que l'employeur ne peut recourir, en cas de contentieux avec un de ses salariés, qu'à des moyens de preuve obtenus à l'aide de processus portés à la connaissance de ses employés⁶¹. La collecte de données personnelles dans le cadre d'opérations de contrôles ne peut se faire à l'insu de l'employé. Les employés doivent donc être informés des dispositifs et moyens de contrôle des outils à leur disposition mis en place par l'employeur. À défaut, la responsabilité de l'entreprise est engagée pour manquement à l'obligation de transparence.

Principe de proportionnalité : la jurisprudence insiste sur le fait que le contrôle de l'employeur sur la messagerie de ses salariés doit être proportionné à la finalité recherchée.

56 PREVOT M., *Op. cit.*, p. 28-32.

57 Soc. 21 décembre 2006. FERAL-SCHUHL C., *Op. cit.*

58 Commission nationale de l'informatique et des libertés (CNIL). « Fiche n°6 : contrôle de l'utilisation d'internet et de la messagerie » dans *Guide pour les employeurs et les salariés*. Paris, 2010, p. 18-19.

59 FERAL-SCHUHL C., *Op. cit.*

60 *Ibid.*

61 Soc. 20 novembre 1991

L'employeur ne peut ainsi exercer un contrôle que lorsqu'il est confronté à un comportement suspect de la part de son employé⁶².

Règlement intérieur ou charte d'utilisation des équipements informatiques

L'information des salariés peut se faire dans le cadre d'un règlement intérieur ou d'une charte annexée au règlement intérieur. Même si le déséquilibre entre employeurs et employés est souvent manifeste dans ces documents, la CNIL soutient leur création lorsque ceux-ci ont pour objectif de sensibiliser les employés aux conditions de sécurité, d'appeler leur attention sur certains comportements de nature à porter atteinte à l'intérêt collectif de l'entreprise⁶³.

Ces documents doivent porter à la connaissance des salariés :

- ♣ le **principe retenu pour différencier les messages électroniques professionnels des messages électroniques personnels**. À défaut d'une identification comme personnels, les messages seront considérés comme professionnels et l'employeur pourra y accéder librement. La nature personnelle d'un message pourra figurer dans l'objet du message ou dans le nom du répertoire dans lequel il est stocké ;
- ♣ les **usages interdits**, car pouvant engendrer une dégradation des services internes ou externes à l'entreprise, comme l'ouverture de fichiers exécutables⁶⁴ ;
- ♣ les **règles de diffusion des ressources** contenues dans le système d'information de l'entreprise, de même que les modalités de cryptage et de décryptage des données confidentielles ou des mots de passe ;
- ♣ les **conditions dans lesquelles s'exerce le pouvoir de contrôle et de consultation par l'employeur** des données produites ou reçues par les salariés⁶⁵ ;
- ♣ les **modalités de fermeture du compte de messagerie**. Il est recommandé à l'employeur d'avertir le salarié de la date de fermeture de son compte afin que ce dernier puisse vider son espace privé.

La charte doit être remise aux salariés sous une forme qui ne prête pas à confusion. L'intégration dans une note de service permet de rendre cette charte opposable, sous réserve d'affichage. Elle peut également être annexée au règlement intérieur. L'employeur doit viser la charte dans le contrat de travail. Il peut également la faire contresigner au moment de la remise des codes d'accès et des identifiants. La charte est ainsi contraignante pour l'employé⁶⁶.

Le rôle-clé des instances de consultation du personnel

Les instances de consultation du personnel, comité d'entreprise dans le secteur privé et comité technique dans le secteur public, sont au cœur du processus d'information et de consultation du personnel.

62 FERAL-SCHUHL C., *Op. cit.*

63 *Ibid.*

64 FAUCHOUX Vincent, DEPREEZ Pierre. *Le droit de l'internet. Lois, contrats et usages*. Paris : LITEC, 2008.

65 Soc. 26 juin 2012

66 Soc. 21 décembre 2006

Pour les entreprises et les établissements publics à caractère industriel et commercial, le code du travail précise dans son article L2323-13 que « le comité d'entreprise est informé et consulté, préalablement à tout projet important d'introduction de nouvelles technologies, lorsque celles-ci sont susceptibles d'avoir des conséquences sur l'emploi, la qualification, la rémunération, la formation ou les conditions de travail. Les membres du comité reçoivent, un mois avant la réunion, des éléments d'information sur ces projets et leurs conséquences sur chacun des sujets mentionnés au premier alinéa ». L'information doit être précise et écrite. L'avis du CE n'est que consultatif⁶⁷.

Pour l'État et les établissements publics administratifs, le décret n°2011-184 du 15 février 2011 précise de son côté, dans son article 34, que « les comités techniques sont consultés, dans les conditions et les limites précisées pour chaque catégorie de comité par les articles 35 et 36 sur les questions et projets de textes relatifs : [...] 4° Aux évolutions technologiques et de méthodes de travail des administrations, établissements ou services et à leur incidence sur les personnels ».

Par extension, la CNIL recommande également que les modalités d'accès de l'employeur aux données stockées sur l'environnement informatique d'un employé absent soient préalablement définies en concertation avec le comité d'entreprise ou le comité technique, afin de garantir le principe de transparence. Cette information doit porter sur les modalités du droit d'accès des salariés aux données et sur la finalité des mesures de contrôle⁶⁸. L'absence de consultation du comité d'entreprise est sanctionnée⁶⁹.

Quelle approche pour le processus d'archivage ?

L'archivage doit concilier l'obligation reconnue aux salariés par la jurisprudence de restituer à leur employeur les éléments matériels et de lui communiquer les informations qu'il détient et qui sont nécessaires à la poursuite de l'activité de l'entreprise (voir ci-dessus) avec le droit des salariés au respect de leur vie privée.

Il semble vraisemblable que les opérations d'archivage doivent se fonder sur les mêmes principes de transparence et de proportionnalité que les opérations de contrôle. D'après la note d'information de la direction des archives de France de 2009, l'employeur doit respecter au moment de l'archivage les six principes suivants :

- ♣ la transparence (obligation d'informer les agents du système d'archivage en place, notamment s'il est automatisé) ;
- ♣ la finalité (conformité avec la législation sur les archives) ;
- ♣ la proportionnalité (un tri doit être effectué dans la conservation des courriers électroniques) ;
- ♣ la sécurité ;
- ♣ la confidentialité (les données ne peuvent être communiquées sans justification précise) ;
- ♣ le droit à l'accès et à la modification (toute personne peut avoir accès aux archives la concernant et peut également corriger, modifier ou supprimer toute information erronée ou fautive la concernant)⁷⁰.

67 Commission nationale de l'informatique et des libertés (CNIL). « Fiche n°6 : contrôle de l'utilisation d'internet et de la messagerie » dans *Guide pour les employeurs et les salariés*. Paris, 2010, p. 18-19.

68 *Ibid.*. Cette obligation est couramment rappelée par la Cour de cassation (Soc. 14 mars 2000).

69 Soc. 7 juin 2006.

70 DIRECTION DES ARCHIVES DE FRANCE. Instruction DITN/RES/2009/007 du 3 juin 2009, consultable en

La CNIL recommande ainsi aux entreprises d'informer leurs salariés des modalités de collecte de leurs messages électroniques, des durées de conservation et des modalités permettant aux salariés d'exercer leur droit d'accès et de rectification⁷¹. Cette approche volontariste d'information des salariés a été retenue par le ministère de la Culture et de la Communication lors de l'archivage du cabinet de Frédéric Mitterrand en mai 2012 : une note de service a été diffusée par la directrice de cabinet du ministre pour demander d'archiver le contenu de leurs messageries professionnelles après avoir expurgé celles-ci des messages couverts par le respect de la vie privée.

La signature par le salarié d'un protocole acceptant la réalisation des opérations d'archivage, en fonction d'un protocole porté à sa connaissance, constitue une solution alternative, quoiqu'extrêmement lourde à mettre en œuvre⁷². Cette seconde solution a été retenue par le ministère de la Santé lors de l'archivage du cabinet de Xavier Bertrand en mai 2012 : chaque membre du cabinet a signé une lettre individuelle de remise de sa messagerie électronique, attestant la nature professionnelle du contenu et le caractère volontaire de la démarche d'archivage⁷³.

1.3. Une utilisation dépendante de l'organisation et de l'utilisateur

Les messageries électroniques « tendent à devenir le support universel de communication écrite, quel que soit le contenu du message »⁷⁴ et sont utilisées pour de nombreux usages qui dépassent largement le simple échange de courriers. L'appropriation de l'outil et l'usage qui en est fait par les utilisateurs sont fortement liés à leur contexte professionnel, à leur culture et aux possibilités offertes par l'écosystème. Comme l'écrivent Philippe Ballet et Jean-Marie Riestch, « chacun réagit en fonction de ses propres ressentis »⁷⁵.

Le comportement des utilisateurs de messageries électroniques n'a cependant pas fait l'objet d'analyses systématiques par les professionnels des archives⁷⁶.

1.3.1. Un outil largement utilisé mais non sans conséquence sur l'organisation du travail

Les technologies informatiques, et singulièrement celles de la messagerie, ont un impact

ligne à l'adresse suivante : <http://www.archivesdefrance.culture.gouv.fr/static/2822> (consulté le 15 mai 2013) rappelée par PREVOT M., *Op. cit.*, p. 35.

71 Commission nationale de l'informatique et des libertés (CNIL), *Op. cit.*, p. 18-19.

72 PREVOT M. *Op. cit.*, p. 51.

73 ALQUIER Éléonore, BENOIST Marie-Élodie, LHOUMEAU Hélène, « Évaluer, sélectionner et collecter les documents numériques non structurés : l'exemple de la collecte 2012 dans les cabinets ministériels (Affaires sociales) », Présentation effectuée au Forum des archivistes de l'Association des archivistes français (AAF), 22 mars 2013.

74 RATIER Daniel, "Les acteurs de l'entreprise à l'épreuve des changements du système d'information", dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*, Paris : La Documentation française, 2012, p. 61.

75 BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 6.

76 PROM C. J., *Op. cit.*, p. 17.

fort sur les méthodes du travail, notamment la gestion du temps et de l'espace : gestion des délais, gestion de l'urgence, traitement d'une information surabondante, vulnérabilité aux dysfonctionnements techniques⁷⁷ en sont des conséquences.

Une appropriation rapide et durable

La messagerie électronique a connu un développement rapide dans l'univers personnel comme professionnel. Plusieurs facteurs expliquent le succès de ce nouvel outil : le développement d'une masse critique d'utilisateurs, grâce à la démocratisation de l'accès à Internet et à l'équipement croissant des foyers et des organisations en micro-ordinateurs ; le caractère intuitif de l'outil et sa diffusion à la fois dans la sphère privée et dans la sphère professionnelle ; une ergonomie assez simple⁷⁸.

Ce succès ne s'est pas ralenti dans le temps, malgré l'introduction de nouveaux outils de communication. L'évolution des protocoles entre équipements locaux et serveurs est sans doute un facteur décisif dans ce succès. L'accès rendu possible à la messagerie *via* tablettes et *smartphones* a préservé son succès. L'usage de ces équipements alternatifs au PC montre à quel point cette technologie devient chaque jour un peu plus intégrée dans la vie quotidienne⁷⁹.

Dans le monde du travail, on estime que **64 % des salariés français disposaient en 2011 d'un matériel informatique « traditionnel » constitué d'un micro-ordinateur relié à un réseau**. Le pourcentage d'utilisateurs dépend de la taille des entreprises et des catégories socioprofessionnelles. Sans surprise, ce sont les cadres qui sont les mieux équipés (97 % de taux d'équipement en 2005 contre 53 % chez les employés, 30 % chez les ouvriers qualifiés et moins de 20 % chez les ouvriers non qualifiés, d'après l'enquête de la DARES)⁸⁰.

Le taux d'utilisation de la messagerie électronique varie fortement d'une catégorie socioprofessionnelle à une autre, les cadres étant trois fois plus équipés que les employés⁸¹.

Un impact sur les rythmes et le temps de travail

Le temps de travail sur écran tend largement à augmenter. En 2005, d'après l'enquête de la DARES, 62 % des salariés travaillaient plus de 3 heures par jour avec leur matériel informatique. La progression par rapport à la précédente enquête de 1998 était particulièrement spectaculaire chez les cadres, puisque 72 % d'entre eux étaient concernés. La DARES signalait que cet accroissement était en grande partie dû à l'utilisation de la messagerie électronique et des outils de bureautique⁸².

Les messageries électroniques sont fréquemment citées comme conduisant à l'auto-accélération des rythmes de travail. Les salariés, et singulièrement les cadres, en viennent d'eux-

77 LASFARGUE Yves, "Panorama de l'utilisation des TIC par les salariés dans leur travail" dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 100.

78 TRAN S., « Quand les TIC réussissent trop bien dans les organisations : le cas du courrier électronique chez les managers », *Management & Avenir*, 2010, 34, p. 200-203.

79 PROM C. J., *Op. cit.*, p. 3.

80 LASFARGUE Y., *Op. cit.*, p. 67-71 et 82-84.

81 *Ibid.*, p. 72-73.

82 *Ibid.*, p. 75-76

mêmes à mettre au point des pratiques de réponse instantanée aux messages reçus ce qui conduit souvent, sans demande explicite, à des pratiques d'emploi permanent de l'ordinateur ou du téléphone mobile, afin d'être connecté en permanence et de ne pas rater de message⁸³.

Le risque de stress et le sentiment de débordement augmentent en conséquence, allant jusqu'à la perte de repères spatio-temporels et à la diminution paradoxale de la productivité. Ils sont particulièrement sensibles pour les utilisateurs de la messagerie électronique : interruption et dispersion, travail dans l'urgence, culture de l'immédiateté, hiérarchisation des tâches selon l'urgence et non selon l'importance. La multi-activité permise par une mobilité croissante (consultation des messageries pendant une réunion) tend à dégrader la productivité des salariés et à accroître les risques de stress⁸⁴.

Une tendance à la surcharge informationnelle

La messagerie a considérablement amplifié le phénomène de surcharge informationnelle. Saturée sous l'effet de l'accumulation des messages, elle nécessite un traitement chronophage, ne permettant pas de hiérarchiser les priorités et ne constituant pas toujours un mode de communication adapté à la densité des échanges. Il est tellement facile d'envoyer des messages que les salariés en abusent, multipliant les échanges et les interlocuteurs mis en copie⁸⁵.

La maîtrise du volume des informations traitées par chaque salarié constitue d'ailleurs l'une des préoccupations majeures des entreprises, tant pour des raisons techniques qu'économiques et sociales. En 2005, d'après l'enquête de la DARES, 63 % des cadres déclaraient recevoir plus de 10 messages par jour, contre 36 % des professions intermédiaires, 34 % des employés et 19 % des ouvriers. L'Observatoire sur la responsabilité sociétale des entreprises (OBSE) estimait en 2011 que 38 % des salariés recevaient plus de 100 messages par jour et que 70 % d'entre eux estimaient souffrir de surcharge informationnelle⁸⁶. Ces chiffres sont vraisemblablement plus élevés aujourd'hui, comme le montrent les enquêtes internes réalisées par les entreprises Canon et Atos Origin. Ces entreprises ont d'ailleurs décidé de prendre des mesures destinées à réduire le volume de messages échangés, soit par maîtrise de l'utilisation de la messagerie (Canon), soit par sa suppression pure et simple en interne et son remplacement par un réseau social d'entreprise (Atos Origin, ministère des Affaires étrangères avec le projet Diplomatie)⁸⁷. Selon le cabinet Gartner, les réseaux sociaux d'entreprises remplaceront le message électronique pour 20 % des communications interpersonnelles à l'horizon 2014⁸⁸.

83 CHEVALLET Romain et MOATTY Frédéric, "Impacts des TIC sur les rythmes, l'autonomie et le contrôle du travail" dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 109 et BENEDETTO-MEYER Marie et KLEIN Tristan. « Le rôle des TIC dans les reconfigurations des espaces et des temporalités » dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 188.

84 CHEVALLET R. et MOATTY F., *art. cit.*, p. 110-112.

85 ROSANVALLON Jérémie, « Les effets des TIC sur l'isolement au travail et les échanges entre collègues » dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 167.

86 OBSERVATOIRE SUR LA RESPONSABILITE SOCIETALE DES ENTREPRISES (ORSE). *Charte pour un meilleur usage de la messagerie électronique dans les entreprises*. Paris : ORSE, 2011

87 LASFARGUE Y., *art. cit.*, p. 78-79. Yves Lasfargue estime que les chartes d'utilisation des outils informatiques prennent de plus en plus en compte cette problématique : ex. OBSERVATOIRE SUR LA RESPONSABILITE SOCIETALE DES ENTREPRISES (ORSE). *Charte pour un meilleur usage de la messagerie électronique dans les entreprises*. Paris : ORSE, 2011

88 RATIER Daniel et SASSON Michel, "Comment pourrait évoluer l'impact des TIC sur les conditions de travail dans

Des frontières brouillées entre vie personnelle et vie professionnelle, entre domicile et lieu de travail

Les messageries électroniques jouent un rôle non négligeable dans le brouillage des frontières spatio-temporelles, notamment en occasionnant une **confusion croissante entre vie privée et vie professionnelle**⁸⁹.

Le phénomène le plus important de ces dernières années est sans doute **l'utilisation croissante des matériels informatiques personnels pour un usage professionnel**, notamment en dehors des heures et du lieu de travail. En 2005, l'utilisation du matériel informatique du domicile pour un usage professionnel concernait 12 % des salariés mais déjà 37 % des cadres⁹⁰. La diffusion des *smartphones* et des tablettes, permettant une connexion croissante au réseau à distance, que ce soit *via* l'accès à un *webmail* ou *via* la connexion sécurisée au poste de travail professionnel a sans doute entraîné une forte croissance de ce pourcentage, ce que confirmera sans doute l'étude de la DARES en cours. En 2011, 70 % des cadres disposaient d'outils de mobilité, contre seulement 26 % chez les non cadres⁹¹. Ce phénomène conduit à une « déterritorialisation » et à une « déspatialisation » des salariés, contribuant à un phénomène de « débordement » le soir, les fins de semaine voire pendant les congés⁹².

L'usage des technologies sur le lieu de travail pour des raisons non professionnelles s'est également développé, notamment pour internet et les messageries électroniques. En 2010, un salarié passait en moyenne 94 minutes par jour sur internet, dont seulement 35 pour des motifs professionnels⁹³. L'intrusion des *smartphones* et des tablettes personnels sur le lieu de travail tend également à accroître le temps de connexion non professionnel à Internet⁹⁴. La messagerie électronique professionnelle est par exemple utilisée pour gérer les problèmes de la vie quotidienne entre conjoints, pour alimenter ses relations avec son cercle d'amis, pour récupérer des informations d'ordre culturel, pour rechercher du soutien parmi ses proches dans le cadre de situations professionnellement tendues⁹⁵.

Ces phénomènes amplifient l'éclatement géographique des activités et de l'information produite et reçue.

les prochaines années" dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 254.

89 BENEDETTO-MEYER Marie et KLEIN Tristan., « Le rôle des TIC dans les reconfigurations des espaces et des temporalités » dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 175.

90 LASFARGUE Y., *art. cit.*, p. 76-77.

91 *Ibid.*, p. 80.

92 BENEDETTO-MEYER M. et KLEIN T., *art. cit.*, p. 175.

93 LASFARGUE Y., *art.cit.*, p. 77-78.

94 RATIER Daniel et SASSON Michel, « Comment pourrait évoluer l'impact des TIC sur les conditions de travail dans les prochaines années », dans CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris, 2012, p. 249-250.

95 LE DOUARIN Laurence, « Gérer la vie quotidienne, occuper des postes d'encadrement : les usages « non professionnels » des TIC par les femmes et leur conjoint sur le lieu de travail », *Politiques et Management public*, avril-juin 2011, vol. 28-2, p. 245-246

*Les cadres, premiers touchés*⁹⁶

Les cadres sont les plus grands utilisateurs des technologies de l'information et de la communication (TIC) et sont ceux dont la messagerie a sans doute le plus modifié les conditions de travail⁹⁷.

La messagerie constitue désormais leur support de travail privilégié ainsi que le principal régulateur de leur temps de travail. En moyenne, un cadre passe 2 heures et 15 minutes par jour à lire et traiter les messages qu'il reçoit.

Plusieurs facteurs ont contribué au succès de la messagerie dans cette population :

- ♣ son caractère plastique : son usage est adaptable aux différentes activités conduites par les cadres ;
- ♣ son caractère asynchrone : son usage évite d'être dépendant de la disponibilité des destinataires au moment du transfert. La transmission par papier n'est plus mobilisée fréquemment que par 32 % des cadres, contre 82 % pour la messagerie ;
- ♣ son caractère de plus en plus synchrone : il lui permet de se substituer à d'autres médias. Pour les discussions, la messagerie est autant mobilisée que le face-à-face, mais bien plus que le téléphone ;
- ♣ son rôle de mémoire organisationnelle opérant une fonction de traçabilité des échanges ;
- ♣ son rôle dans un réseau d'acteurs dans lequel les jeux d'influences ont une importance stratégique.

Les cadres se sentent les premières victimes du phénomène de surcharge informationnelle et ont l'impression de se laisser déborder par l'outil, tout en ayant l'impression de faire le maximum pour endiguer le flux. Cette surcharge génère de nouvelles tâches, comme la priorisation des messages eux-mêmes, ainsi que la lecture et le transfert des messages inappropriés, tâches auparavant effectuées en amont par les secrétariats. Les messages sont toujours plus volumineux et multi-destinataires, saturant les réseaux et entraînant des échanges à géométrie variable. La prise de décision suit désormais deux chemins parallèles, papier et électronique.

L'usage de la messagerie induit cependant un autre rapport au temps chez les cadres. La réactivité est désormais placée au cœur de leurs critères distinctifs, puisque certains subissent une véritable pression temporelle du court terme. Les cadres consacrent ainsi de plus en plus de temps à traiter en temps réel des informations aux caractères épars et hétérogène, ce qui génère une forte tension entre réactivité et flexibilité, entre urgence et importance. Le délai tampon de réponse permis par le papier n'existe plus. La messagerie autorise également les cadres à s'affranchir du cadre temporel et spatial. 73 % d'entre eux déclarent travailler en dehors des heures de bureau : 66 % dans la fonction publique d'État⁹⁸, 58 % le soir, 50 % le week-end et 38 % pendant les vacances⁹⁹. Le travail en débordement est très répandu dans cette population particulièrement

96 BRETESCHE Sophie, CORBIERE François de, GEFROY Bénédicte. « La messagerie électronique, principal métronome des activités de cadre », *La Nouvelle Revue du Travail*, 2012, 1 (consultable en ligne sur revues.org) et CIHUELO Jérôme, SAINTIVE Brigitte, « De l'émiettement du travail intellectuel : déstructuration et dépossession de l'activité de bureau par la messagerie électronique », *Communication & Organisation*, 2000, n° 17 (accessible sur revues.org).

97 LASFARGUE Y., *art. cit.*, p. 97.

98 CONSEIL D'ANALYSE STRATEGIQUE, « Quel est l'impact des TIC sur les conditions de travail dans la fonction publique ? », *La note d'analyse*, janvier 2013, n° 318, p. 4.

99 Sondage OpinionWay 2011 cité par PRAS Bernard. « Entreprise et vie privée. Le 'privacy paradox' et comment le

concernée par la question de l'équilibre entre vie professionnelle et vie privée. L'arrivée de la connexion permanente démultiplie les risques d'usages intensifs pour la santé et la question du droit à la déconnexion pourrait revenir en force¹⁰⁰.

La messagerie constitue donc une source de stress au travail, notamment du fait de la pression temporelle. Répondre rapidement aux messages devient une preuve d'engagement et induit donc un traitement en continu. L'impossibilité de se connecter alimente également un sentiment d'inquiétude. Les cadres ont l'impression d'un émiettement croissant de leur activité et d'une pollution croissante de la vie et des espaces personnels par leur travail, le traitement sur le fond des dossiers ne pouvant être réalisé que le soir ou le week-end, dans des espaces temps déconnectés du flux informationnel.

L'impact des messageries sur les méthodes de travail des cadres se répercute naturellement sur la gestion du cycle de vie de cette masse d'information. Faute de secrétariat ou d'assistants, les cadres ont tendance à effectuer des opérations de tri et de classement imprécis et se contentent des fonctionnalités dites d'archivage fournies par leur client de messagerie. Ces opérations sont souvent rendues indispensables par le paramétrage de leur espace de stockage (export des messages dans une boîte mensuelle), mais ces exports rassurent plus qu'ils ne servent. Ce système ne reste possible et viable que dans la mesure où les utilisateurs disposent d'un stockage de grande capacité et de fonctions de recherche performantes¹⁰¹.

1.3.2. Les facteurs qui conditionnent l'utilisation de la messagerie

Plusieurs facteurs influencent la manière dont un individu utilise la ou les messageries électroniques qui sont mises à sa disposition, dans sa vie personnelle comme dans sa vie professionnelle : des facteurs techniques (paramétrage de l'outil messagerie et place de la messagerie dans le système d'information de l'organisation) ; des facteurs fonctionnels (finalité donnée par l'organisation à l'outil messagerie) ; des facteurs comportementaux (manière de gérer ses documents). Un même outil sera donc utilisé de multiples manières au sein d'une même organisation, ce qui rend encore plus difficile la mise en place d'une politique d'archivage.

Facteurs techniques

Il est rare qu'un utilisateur dispose de l'administration complète des outils informatiques à sa disposition. Dans une grande majorité des cas, que ce soit dans sa vie personnelle (utilisation de services de messageries offerts en ligne) comme dans sa vie professionnelle (utilisation d'un service de messagerie fourni par l'entreprise), le(s) outil(s) informatique(s) qu'il utilise ont déjà fait l'objet d'un paramétrage, que ce soit par le fournisseur des matériels, des logiciels ou des services. Ces paramétrages techniques influencent considérablement, même parfois inconsciemment, la manière dont il utilise sa ou ses messageries.

dépasser », *Revue française de gestion*, 2012, n° 224, p. 87-94.

100 RATIER D. et SASSON M., *art. cit.*, p. 261-262.

101 CIHUELO Jérôme, SAINTIVE Brigitte, « De l'émiettement du travail intellectuel : destructuration et dépossession de l'activité de bureau par la messagerie électronique », *Communication & Organisation*, 2000, n° 17 (consultable sur revues.org).

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

Le tableau ci-dessous s'efforce de recenser les cas possibles et leur impact sur les usages¹⁰² :

Catégorie de paramètre	Possibilités de paramétrage	Impact sur l'usage
Mode de gouvernance du système	- modalités de création des comptes ; - modalités de désactivation des comptes.	- nombre d'adresses et de comptes disponibles par individu ; - disponibilité du compte après le départ du salarié par celui-ci et par l'entreprise.
Degré d'ouverture du système	- utilisation exclusivement interne <i>via</i> le réseau privé (intranet) ; - utilisation ouverte vers l'extérieur <i>via</i> un réseau public (internet)	- nombre de correspondants ; - accès à distance aux différents outils mis à disposition par l'employeur (messagerie, GED, ERMS) ; - transfert des messages sur le(s) compte(s) de messagerie personnel(s).
Environnement technique	- utilisation exclusive pour la communication, la gestion et les échanges de documents et d'informations ; - existence d'autres outils de communication, de gestion et d'échanges de documents et d'information : téléphone, messagerie instantanée, webcam, serveur de transfert de fichiers, GED, ERMS	- concentration des informations ou répartition entre un ou plusieurs outils ; - utilisation de la messagerie pour les usages de communication synchrone et de gestion des documents.
Protocole d'accès	- protocole POP3 - protocole IMAP - protocole HTTP	- conservation systématique ou non d'une trace des messages sur le serveur ; - accès synchronisé au compte avec différents équipements (équipements professionnels/personnels ; équipements fixes/équipements mobiles).
Choix et contraintes en matière de stockage	- gestion des pourriels et des listes de diffusion au niveau du serveur de réception ; - définition de quotas de stockage pour chaque adresse sur les espaces serveurs, avec ou sans règles de purge automatiques ¹⁰³ ; - capacité illimitée de stockage sur les espaces serveurs, si besoin avec mise en place de solutions adaptées de stockage (type <i>cloud computing</i>) ; - mise à disposition des utilisateurs de supports amovibles de stockage ou d'espaces de stockage sur des serveurs de fichiers partagés	- répartition des messages entre espaces serveur et espaces locaux sur disque dur (avec risques que cela comporte : absence de sauvegarde, absence d'accès à distance) ; - nombre, fréquence et caractère plus ou moins automatique des éliminations ; - recours aux fonctionnalités « d'archivage » offertes par les clients de messagerie ¹⁰⁴
Choix et contraintes en matière d'accès	- accès <i>via</i> un client local ; - accès <i>via</i> une <i>webmail</i> .	- répartition des messages entre espaces serveur et espaces locaux sur disque dur (besoin d'accès à distance).
Fonctionnalités du logiciel	- existence d'un système de	- possibilité d'organiser un système de

102 Tableau établi en partie à partir de BIA FIGUEIREDO Maria et KALIKA Michel. *La communication électronique*. Paris, 2009, p. 18 et de PROM C. J., *Op. cit.*, p. 9-10.

103 Les services informatiques peuvent avoir tendance à brider les capacités de stockage à disposition des utilisateurs en imposant une taille limitée à la boîte aux lettres. Cf. BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 6.

104 *Ibid.*, p. 9.

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

retenu (<i>webmail</i> et client) : classement et capacité de créer des dossiers locaux	classement par défaut (boîte de réception/boîte d'envoi/brouillons) ; - organisation des dossiers locaux (souvent alphanumérique strict)	classement perfectionné ; - dépendance pour la restitution du système de visualisation permis par l'outil.
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : impression et export	- export unitaire vs. Export global ; - possibilité d'imprimer en PDF ; - type de format d'export proposé par le client : le plus répandu est le format pst (Outlook) ¹⁰⁵ .	

Ces paramètres techniques vont donc avoir des conséquences sur le degré de responsabilité des utilisateurs dans la gestion de leur messagerie (obligation ou non de mettre en œuvre des opérations régulières de gestion de leur messagerie)¹⁰⁶, sur le lieu de stockage des différents messages (équipement personnel/équipement professionnel, serveur/disque dur ou espace sur un serveur de fichier partagé, supports amovibles), sur le nombre de copies disponibles de chaque message, sur les formats dans lesquels ceux-ci sont stockés, etc¹⁰⁷.

Facteurs fonctionnels

Une institution peut utiliser le même outil technique de messagerie avec différentes finalités, ce qui implique des paramétrages, notamment en matière d'accès, différents.

On distingue traditionnellement trois grandes catégories d'usages fonctionnels des messageries, en fonction du degré de personnalisation et de sécurisation des échanges autorisés aux utilisateurs et de la nature des échanges que l'on souhaite voir réalisés avec l'outil :

- ♣ **les messageries interpersonnelles**, les plus répandues, permettent à des individus disposant chacun d'une adresse électronique nominative, d'échanger des messages, à l'intérieur ou à l'extérieur de l'organisation ;
- ♣ **les messageries fonctionnelles**, plus génériques, permettent à un individu de contacter une structure matérielle ou immatérielle (service, association) ou une fonction. Les adresses électroniques ne sont pas la propriété d'un individu et peuvent donc être transmises à de nouveaux utilisateurs en cas de changement d'organisation ou de personnel. Les plus connues permettent à des individus de poser des questions très générales à des organisations et sont les plus répandues (ex. : contact@organisation). Ces messageries sont souvent administrées par un groupe d'utilisateurs ;
- ♣ **les messageries d'autorité/messageries officielles** qui permettent à des entités de se contacter et de se transmettre des ordres ou des informations. Les informations échangées par le biais de ces messageries nécessitent souvent des signes de validation hiérarchique forts et n'appellent pas forcément de réponse. L'utilisation de l'outil est généralement très encadrée (droits d'accès, circuit de validation, formalisme des échanges). Ces messageries officielles sont plus particulièrement utilisées dans des institutions à l'organisation très hiérarchisée, comme le ministère de la Défense (ex. : messageries d'autorité Melind@ et SIC21 de l'état-major de la Marine nationale). Même si elle ne s'appuie pas sur des outils de messagerie électronique fournis par le marché, la correspondance diplomatique gérée par le

105 PROM C. J. *Op. cit.*, p. 22-23.

106 *Ibid.*, p. 16-17.

107 *Ibid.*, p. 10.

ministère des Affaires étrangères peut être assimilée à une messagerie d'autorité.

À chaque catégorie de messagerie va répondre des types d'utilisateurs et des types d'usages différents, que ce soit au niveau de la forme comme du fond des messages échangés.

Facteurs comportementaux

Les utilisateurs sont des acteurs qui, au travers du processus d'appropriation et des contraintes, notamment techniques, auxquels ils sont soumis, font évoluer les messageries électroniques à partir de leurs usages sur des fonctions pour lesquelles elles n'avaient pas été conçues au départ. La relation de l'utilisateur à la messagerie électronique est dynamique et se modifie dans le temps, ce qui peut expliquer que le comportement des usagers paraisse en décalage par rapport à l'emploi préconisé au départ. Les études montrent un enrichissement et une diversification de l'usage de l'outil au fur et à mesure du temps.

Alors que la messagerie avait été conçue pour assurer une communication asynchrone, sur le modèle exact de la correspondance papier, et était considérée en conséquence comme un outil pauvre, son usage a été progressivement détourné par les utilisateurs. Trois nouvelles fonctionnalités ont été conçues, rendues possibles par l'appropriation de l'outil et par les innovations technologiques. Elles ont conféré à la messagerie des propriétés proches d'autres technologies :

- ♣ **la communication synchrone** : cet usage a été rendu possible par l'accroissement du débit des réseaux et la baisse des coûts d'abonnement. Il a encore été amplifié par le développement des *smartphones*. La messagerie électronique s'est substituée au téléphone fixe et a permis de prolonger les échanges en dehors des heures de travail ;
- ♣ **la coordination des tâches et des projets** : la messagerie électronique tend à devenir un outil de travail collaboratif. Plusieurs facteurs ont contribué au développement de cet usage : capacité de stockage accrue ; accès en situation de mobilité à partir d'équipements multiples (personnels comme professionnels) ; équipement massif des utilisateurs avec des messageries, alors que les outils dédiés au travail collaboratif sont moins répandus et moins facilement accessibles en situation de mobilité. La messagerie électronique est devenue une véritable porte d'entrée dans les technologies au point que certains éditeurs ont considéré qu'elle devait devenir le point d'entrée des outils de gestion de contenu ;
- ♣ **la mémorisation individuelle ou organisationnelle** : l'outil est utilisé pour la gestion de l'information (veille, *records management*) et est devenu un outil de mémorisation et de codification des informations, devenant une mémoire institutionnelle et personnelle. La messagerie électronique sert de plus en plus à la sauvegarde des données, messages et pièces jointes, grâce à l'augmentation des capacités de stockage – c'est particulièrement le cas des messageries accessibles depuis des *webmails*, appuyées sur un stockage en mode *cloud computing*. Par ailleurs, les messageries électroniques permettent de saisir et de transmettre rapidement des documents, notes ou comptes rendus de réunions, brouillant ainsi davantage les fonctions des autres outils mis à disposition des utilisateurs, notamment les logiciels de traitement de texte¹⁰⁸.

La possibilité d'accéder à distance aux messages et à l'information, notamment en dehors du lieu et des horaires de travail, est sans doute aujourd'hui le facteur qui influence le plus sur l'usage

108 TRAN S., *art. cit.*, p. 209-211.

constaté des messageries électroniques. À défaut d'une *webmail* disposant d'un espace de stockage suffisant, 77 % des utilisateurs transfèrent des messages professionnels sur leurs comptes personnels. De nombreuses personnes utilisent uniquement des comptes de messagerie accessibles *via* une *webmail* comme une stratégie de stockage et de sauvegarde de leurs documents importants¹⁰⁹.

Les pratiques de gestion des messageries sont éminemment individuelles et dépendent à la fois du temps dont chaque utilisateur dispose, de son appétence pour les opérations de gestion de l'information (tri, classement, élimination) et de l'affectivité plus ou moins grande qu'il apporte à cette information. De la complète négligence à une véritable utilisation en matière de gestion de l'information, toutes les attitudes sont possibles¹¹⁰.

L'hétérogénéité des pratiques et des comportements observés dans l'usage de la messagerie est en grande partie liée aux modalités de déploiement, de formation et de communication de l'outil. Définies dans la grande majorité des cas par les seuls services informatiques, ces modalités se concentrent sur l'appropriation technique de l'outil et ne donnent aux utilisateurs aucune consigne d'ordre managérial ou organisationnel. Les utilisateurs ne sont donc accompagnés dans leur appropriation et leur mode d'utilisation de l'outil que par ces seules directives techniques¹¹¹.

Les études de cas effectuées par les membres de l'équipe VITAM dans chacun de leurs ministères respectifs montrent à quel point l'environnement technique influe sur l'utilisation des messageries et, partant, sur la localisation et la collecte des archives concernées (voir annexe 3).

Conclusion : une grande complexité pour l'archivage

Simple d'utilisation, largement répandues sur les équipements informatiques personnels comme professionnels, les messageries électroniques sont devenues des outils indispensables de création, d'expédition, de réception, de stockage et de recherche de l'information. **En raison de son succès auprès des populations de cadres, l'information stratégique pour tout individu et pour toute organisation y est de plus en plus concentrée. En lieu et place des notes administratives, ce sont désormais des messages qui véhiculent la prise de décisions, à tous les niveaux des organisations. L'archivage des informations qui sont transmises par cet outil est donc de première nécessité.**

L'archivage des messageries fonctionnelles et des messageries officielles/d'autorité ne pose aucune difficulté autre que technique. En effet, leur finalité claire, leur usage encadré et leur formalisme accentué ne posent guère de problème, notamment juridique.

L'archivage des messageries interpersonnelles se révèle en revanche extrêmement problématique, étant donné la diversité des usages et les contraintes légales liées à la protection de la vie privée qui pèsent sur leur contenu.

Les archivistes abordant la conservation des messages électroniques de valeur historique sur

109 PROM C. J., *Op. cit.*, p. 4.

110 *Ibid.*, p. 18.

111 TRAN S., *art. cit.*, p. 208-209.

le long terme sont alors confrontés à plusieurs défis :

- Les formats de stockage, les systèmes de stockage ;
- Les politiques institutionnelles et les pratiques individuelles de gestion sont fortement malléables ;
- La gestion des messageries électroniques repose sur des individus et des entités qui n'ont aucun intérêt à la conservation à long terme (informaticiens, fournisseurs de comptes de messagerie voire utilisateurs eux-mêmes) ;
- Les exigences légales, les contraintes financières et des pratiques individuelles relâchées de gestion de l'information conduisent nombre d'institutions et d'individus à négliger ou à détruire activement les messages électroniques ;
- La liberté, l'ubiquité et la banalité offertes par la gestion des messageries, du point de vue de l'utilisateur font que la gestion des messageries est rarement une priorité, et encore moins leur préservation.

Cela ne signifie pas que la conservation des messages électroniques, du moins d'un point de vue juridique comme technique, est impossible ou même difficile¹¹². C'est souvent dans le repérage de l'information pertinente à archiver et dans sa sélection au milieu de ce qui s'apparente le plus souvent à un vrac sans organisation interne que résident les véritables difficultés. Tout programme de conservation des messages électroniques doit reposer sur une solide compréhension des préférences et des comportements des utilisateurs.

112 PROM C. J., *Op. cit.*, p. 23.

II - Quelle approche archivistique de l'objet messagerie ?

L'analyse des usages des messageries effectuée, il est indispensable d'établir des règles de gestion et de traitement des messageries, et ce, afin de leur assurer une conservation durable et intègre.

Mais cette approche dite « archivistique » ne va pas de soi. En effet, face à la masse de messageries et de messages, quelles règles de sélection et de tri appliquer ? Quels formats recommander pour l'export ? Comment peut-on redonner accès aux messageries collectées et archivées ?

Le schéma ci-dessous résume l'ensemble des questions qu'un archiviste doit se poser au moment de définir une stratégie d'archivage des messageries de l'organisation dans laquelle il travaille.

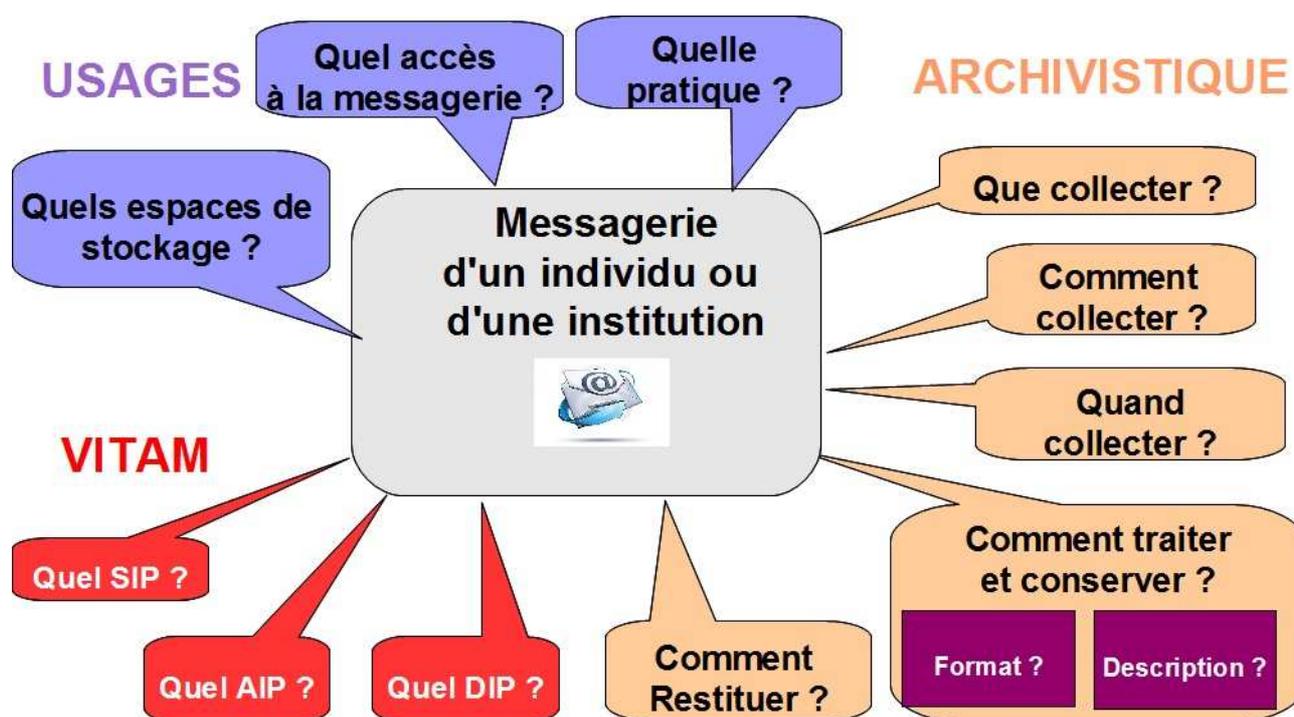


Fig. 3 : les différents aspects de l'archivage d'une messagerie

À cet effet, les programmes et expériences menés par d'autres services d'archives à l'étranger constituent une aide précieuse pour entamer cette réflexion.

2.1. Les réflexions engagées au niveau national et international

Depuis maintenant 15 ans, plusieurs acteurs se sont penchés sur la question de l'archivage des messageries électroniques : services informatiques et autres acteurs du numérique (comme la FedISA en France), associations professionnelles et établissements patrimoniaux. Si les *records managers* et les archivistes ont été leaders sur le sujet, ce sont désormais les services patrimoniaux des bibliothèques universitaires et des universités qui poursuivent principalement la réflexion, notamment en Grande-Bretagne (Bodleian Library) ou aux États-Unis (Smithsonian Institution Archives, Harvard University Library).

Il est possible de regrouper les travaux réalisés sous trois catégories :

- ▲ **l'approche « pédagogique »**, la plus répandue, par lesquels les acteurs du monde du numérique et de l'archivage essaient d'alerter les utilisateurs sur les conséquences de leur usage de l'outil messagerie et qui essaient de leur fournir des conseils et guides de bonnes pratiques ;
- ▲ **l'approche « collecte et conservation »**, avec quelques projets-phares, qui s'efforce de répondre aux problématiques de collecte et de conservation à moyen comme à long terme des messageries électroniques, notamment par la conception d'outils techniques ;
- ▲ **l'approche diplomatique**, fondée principalement sur les questions d'intégrité *via* le renouveau de la diplomatie initié par l'université de Colombie britannique et notamment par le professeur Luciana Duranti dans le cadre du groupe InterPARES.

La synthèse des travaux réalisés qui va suivre se fonde sur la revue de la littérature professionnelle existante et sur des échanges avec plusieurs services d'archives et groupes de réflexions qui ont bien voulu répondre aux interrogations et aux demandes de précisions formulées par les auteurs du présent document. Prolonger les échanges initiés avec certaines institutions et organisations, comme les Archives de l'État en Belgique ou la Wellcome Trust Fondation, par une collaboration plus étroite sur le sujet est d'ailleurs à l'étude.

2.1.1. L'approche « pédagogique »

Nombreux sont les associations professionnelles et les services d'archives, qu'ils se situent au niveau d'un État, d'un territoire ou d'une organisation en tant que telle (dans le secteur public comme dans le secteur privé), à avoir édicté des politiques et des règles qui évoquent la question des messageries électroniques. Certaines choisissent de diffuser des textes dédiés aux messageries – on citera ceux des Archives de l'État en Belgique¹¹³, des archives des états d'Alabama, de Caroline du Sud ou du Texas, de Bibliothèque et Archives nationales du Québec (BAnQ)¹¹⁴, des Archives provinciales du Nouveau-Brunswick¹¹⁵, de The National Archives¹¹⁶ ou de l'Organisation des

113 SOYEZ S., *Op. cit.*

114 BIBLIOTHEQUE ET ARCHIVES NATIONALES DU QUEBEC, Orientations pour la gestion documentaire des courriels du gouvernement du Québec, 2009, 20 p., disponible en ligne : http://www.banq.gc.ca/documents/services/archivistique_ged/Orientations_Gestion_courriels_gouv_Quebec_2009-01-09_VF.PDF (consulté le 27 mars 2013).

115 Fiches de procédures et de conseils disponibles en ligne à l'adresse suivante : <http://archives.gnb.ca/Archives/RecMan.aspx?Section=2&culture=en-CA> (consulté le 15 mai 2013)/

Nations Unies (ONU)¹¹⁷ -, d'autres se contentent de modifier ou d'actualiser des politiques existantes, où les messageries ne sont considérées que comme une typologie technique comme une autre – c'est le cas des états de l'Idaho, de l'Oklahoma et de Washington. La première approche est critiquée comme constituant une mauvaise stratégie qui attire l'attention sur une catégorie technique au détriment de la gestion globale des documents d'activité de l'organisation dont les documents véhiculés par messagerie électronique ne constituent qu'une section¹¹⁸.

Cette approche a été naturellement reprise en France, même si le recensement des outils pédagogiques conçus par les services d'archives internes aux organisations est particulièrement difficile. Ces documents, s'ils existent, ne font généralement l'objet que d'une diffusion interne sur intranet et sont rarement portés à la connaissance du public externe. C'est le cas par exemple des guides réalisés par les Archives nationales de France¹¹⁹ et la Bibliothèque nationale de France. En revanche, plusieurs associations ont rédigé et diffusé des livres blancs et des guides à destination de leurs adhérents ou de leur public-cible, en vue de les sensibiliser à l'importance de la gestion et de l'archivage des messageries électroniques : Association des archivistes français (AAF) ; Club des responsables de politiques et projets d'archivage (CR2PA)¹²⁰ ; Fédération de l'ILM du Stockage et de l'Archivage (FedISA)¹²¹. La direction des Archives de France (DAF), de son côté, a publié le 3 juin 2009 une instruction diffusant et commentant la directive publiée par les Archives de l'État en Belgique¹²².

La structuration de ces différents documents est souvent similaire, même si leurs approches sont différentes (gestion des risques, efficacité administrative, conservation du patrimoine) et comprend généralement les sections suivantes :

- ♣ la présentation générale de la problématique et du contexte (notamment juridique) et justification de l'importance du sujet et de la mise en place de procédures ;
- ♣ la répartition des responsabilités entre les différents acteurs et modalités de sensibilisation de ceux-ci (managers, rédacteurs, services informatiques, *records managers* et archivistes) ;
- ♣ les conseils pour la rédaction des messages (quand en créer ? Comment en créer ? Comment les rédiger ?) ;
- ♣ les conseils pour identifier les messages qui doivent intégrer le système d'archivage (identification des *records*¹²³) ;
- ♣ les conseils pour la gestion des messages par les utilisateurs (comment les classer ? Avec quels outils ?) ;
- ♣ la présentation des solutions disponibles pour les organisations, de l'outil de *records management* au tiers archivage, en passant par les solutions de capture et de stockage automatisées disponibles sur le marché ;

Quel que soit leur public-cible (*records managers*, décideurs, services informatiques), tous

116 THE NATIONAL ARCHIVES. *Guidelines on developing a policy for managing email*. Londres, 2004. 44 p.

117 <http://archives.un.org/ARMS/content/managing-emails-records> (consulté le 15 mai 2013).

118 The InterPARES 3 Project, TEAM Italy, *Guidelines and Recommendations for E-Mail Records Management and Long-Term Preservation*, 2011, p. 10.

119 ARCHIVES NATIONALES, *Op. cit.*, 60 p.

120 CLUB DES RESPONSABLES DE POLITIQUE ET PROJETS D'ARCHIVAGE [CR2PA], *Op. cit.*

121 BALLEST P., RIETSCH J.-M., *Op. cit.*.

122 Instruction DITN/RES/2009/007du 3 juin 2009, consultable en ligne à l'adresse suivante : <http://www.archivesdefrance.culture.gouv.fr/static/2822> (lien vérifié le 15 mai 2013).

123 Voir par exemple les documents de l'Organisation des Nations-Unies, des Archives provinciales du Nouveau-Brunswick, de The National Archives.

ces documents insistent sur le **rôle de l'utilisateur final et sur sa responsabilité**¹²⁴.

2.1.2. L'approche « collecte et conservation »

Les travaux de ces dernières années ne se sont cependant pas limités à cette approche pédagogique. Plusieurs travaux méritent à ce titre une attention particulière autant par leur souhait d'appréhender les caractéristiques propres des messageries électroniques, dans leur forme intellectuelle et physique, que pour leur objectif de concevoir des outils à même d'assurer la collecte et la conservation de messages électroniques. Les premiers ont eu une approche unitaire, les suivants une approche plus globale.

Le projet DAVID (*Digitale Archivering in Vlaamse Instellingen en Diensten*) mené de 1999 à 2003 par les Archives municipales d'Anvers a été le premier à proposer un cadre d'analyse de la problématique et à passer en revue les différentes options politiques et techniques possibles (impression, duplication intégrale, export vers un espace réseau et capture dans un outil de *records management*). Le rapport final rédigé par Filip Boudrez expose la manière dont les messages électroniques peuvent être gérés et archivés et comment la municipalité d'Anvers a conçu et développé une procédure de *records management* et d'archivage des messages électroniques gérés par son administration et ses établissements publics.

Ce projet a abouti au développement et à l'implémentation de deux outils :

- ♣ FilingToolbox : module additionnel du client Outlook (capture et saisie des métadonnées, ainsi que classement des messages et de leurs pièces jointes) et extension de Windows Explorer pour l'enrichissement des métadonnées et la réplique de structures de fichiers ;
- ♣ ArchivalToolbox : outil de migration et d'encapsulation au format XML, ainsi que de lecture des *Archival Information Package* (AIP)¹²⁵.

Le projet expérimental mené par les Archives nationales des Pays-Bas et le ministère néerlandais de l'Intérieur et des Affaires royales en 2003 s'est également penché sur la question des messageries électroniques. Il s'est plus particulièrement intéressé à la question de l'intégrité des messages électroniques et des caractéristiques essentielles des messages à conserver pour garantir leur intégrité. Le rapport pèse les avantages et les inconvénients de trois stratégies de conservation : la migration vers un autre logiciel ; la constitution d'enveloppes XML et l'émulation, concluant que l'utilisation du langage XML constituait la stratégie la plus prometteuse. Le projet a conçu un schéma XML ainsi qu'un module complémentaire au client Outlook pour convertir les messages au format XML qui n'est malheureusement plus disponible en ligne¹²⁶.

Ces deux projets se sont focalisés sur la gestion unitaire des messages électroniques.

En décembre 2002, les National Archives of Australia (NAA) ont publié les principes de la politique d'archivage électronique d'Australie¹²⁷, *An approach to the preservation of digital records*.

124 Voir par exemple la présentation disponible sur le site internet des Archives provinciales du Nouveau-Brunswick.

125 BOUDREZ Filip, *Filing and archiving e-mail*. Antwerp, 2006, notamment p. 42.

126 DIGITAL PRESERVATION TESTBED. *Op. cit.* et PROM C. J., *Op. cit.*, p. 6-7.

127 THE NATIONAL ARCHIVES OF AUSTRALIA, *An approach to the preservation of digital records*, 2002, disponible en ligne : http://www.naa.gov.au/images/an-approach-green-paper_tcm2-888.PDF (consulté le 15 mai 2013).

Cette politique a été l'élément fondateur du développement de leur outil de conversion des fichiers et d'extraction de métadonnées nécessaires à leur pérennisation. Depuis peu, cette approche globale s'est combinée au plan de continuité numérique (*Digital Continuity Plan*) qui conduira, d'ici 2015, à la collecte unique d'archives numériques délaissant ainsi les missions de collecte des archives papier¹²⁸.

L'approche des NAA est globale : les messages ne sont pas uniquement analysés comme des objets techniques mais comme des documents qui nécessitent une évaluation diplomatique et documentaire. Pour autant, les NAA ont publié un guide de bonne gestion des messageries¹²⁹. Les recommandations envisagent même le cas des *Voice mail messages*¹³⁰ même si aucune méthodologie d'archivage n'est avancée.

Pour traiter les messages, les NAA ont développé en interne deux outils dont ils assurent la maintenance, revendiquant fortement leur indépendance et la maîtrise des processus de traitement¹³¹ :

- **XENA** (*XML electronic normalising for Archives*) : qui assure le travail de « normalisation » des archives électroniques en XML¹³² pour constituer des « *preservation master* »¹³³
- **DRP** (*Digital Preservation recorder*)¹³⁴ : qui assure le contrôle de tout le *workflow d'ingest* y compris la vérification automatique que les supports de transfert contiennent bien les documents annoncés et que les fichiers ne sont pas corrompus (espace de quarantaine).

Pour autant, ces transferts de messageries ne sont pas automatiques, les outils n'offrant pas cette possibilité.

Le projet Collaborative Electronic Records Project (CERP), mené conjointement par la Smithsonian Institution Archives (SIA) et le Rockefeller Archive Center (RAC) de 2003 à 2006, a, de son côté, abordé la problématique des messages électroniques au niveau global du compte de messagerie. L'objectif était de définir un modèle de collecte conforme aux préconisations de la norme OAIS. Ces deux institutions se sont ensuite associées avec l'E-Mail Collaborative Initiative (EMCAP), regroupant plusieurs services patrimoniaux – North Carolina State Archives, Pennsylvania State Archives, Kentucky Department of Libraries and Archives, en vue de développer un schéma XML capable de gérer un compte de messagerie dans son ensemble (organisation, messages et pièces jointes), tout en préservant les métadonnées au niveau unitaire. En est résulté la conception de l'E-Mail Account XML Schema ainsi que la conception d'un outil de

128 NAA, Digital Continuity plan (<http://www.naa.gov.au/records-management/agency/digital/digital-continuity/plan/index.aspx>) et Digital Transition Policy (<http://www.naa.gov.au/records-management/digital-transition-policy/>)

129 NAA, *Managing emails*, disponible en ligne, 2013 :

<http://www.naa.gov.au/records-management/agency/digital/managingemail/index.aspx>

130 NAA, *FAQ on Records management*, disponible en ligne : <http://www.naa.gov.au/records-management/faqs/index.aspx#section10> (consulté le 15 mai 2013).

131 Ces outils sont uniquement open source. Leur maintenance est exclusivement assurée par les NAA. Cf. HESLOP, Helen, DAVIS Simon, WILSON, Andrew, *An approach to the preservation of digital records*, National Archives of Australia, décembre 2002, p.15. Disponible en ligne : http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161_tcm16-70792.PDF (consulté le 21/05/2013).

132 Les NAA ne souhaitent pas parler de conversion. Elles emploient le terme de « normalisation » qui n'intervient qu'une seule fois dans le processus d'archivage au moment de la constitution de l'AIP.

133 En parallèle, les NAA conservent une version de diffusion appelée « *accession record* » alors que les « *preservation masters* » constituent la version de conservation. Cf. CARDEN David, *Digital Archiving at the National Archives of Australia: Putting Principles into Practice*, intervention au Congrès international des Archives à Brisbane, août 2012.

134 DPR, version 5.1.2, 19 février 2013, téléchargeable en ligne : <http://dpr.sourceforge.net/>

transformation de comptes de messageries mbox en fichiers conformes au schéma XML, le CERP E-Mail Parser¹³⁵.

2.1.3. L'approche « diplomatique »

Les travaux les plus récents adoptent une approche différente de la question des messageries électroniques, en se focalisant sur les problématiques d'authenticité et d'intégrité, examinées au prisme de la diplomatique.

Le projet anglais InSPECT, subventionné par JISC et mené de 2008 à 2010 par le King's College de Londres et The National Archives, s'est ainsi interrogé sur la notion de données essentielles (*significant properties*) appliquée aux messages électroniques, avec l'objectif de définir les stratégies à mettre en œuvre, lors des opérations d'archivage et de pérennisation, pour garantir la préservation de ces informations et de leur intégrité. 64 données ont été considérées comme essentielles pour garantir l'authenticité et l'intégrité d'un message : 14 dans les en-têtes et 50 dans le corps lui-même (cf. infra, section 2.2.1). Le projet s'est également intéressé aux conséquences des opérations de migrations sur la conservation de ces données. Une série de tests a donc été réalisée à partir de jeux de données fournis par les membres du groupement pour comparer les extractions de données réalisées sur des messages au format natif et sur des messages convertis dans des formats cibles (XML, mbox, txt)¹³⁶.

Le groupe InterPARES, dans le cadre de la 3^e phase du projet, s'est penché sur les problématiques d'authenticité dans les processus d'archivage des messages électroniques. Le rapport rédigé par l'équipe italienne insiste, au regard des critères fondant l'authenticité des documents électroniques élaborés lors des précédentes phases du projet, sur la nécessité d'appliquer un traitement adapté aux messages électroniques et édicte une série de recommandations à ce sujet¹³⁷.

2.2. Bilan des expériences

À partir des initiatives analysées, il est nécessaire de dresser un état des lieux de la réflexion archivistique et d'en souligner ces limites ou les futures expérimentations à mener pour répondre aux besoins du contexte français.

135 FERRANTE Riccardo, SCHMITZ FUHRIG Lynda. *Digital Preservation : Unsing the Email Account XML Schema*, Washington : Smithsonian Institution, [2007], 8 p.

136 KNIGHT J. *InSPECT Project Document - Significant Properties Testing Report : Electronic Mail*. London : JISC/The National Archives/King's College London, 2009.

137 Les 5 caractéristiques constituant un document électronique ayant valeur de record sont les suivantes : fixité ; contextualisation ; lien archivistique ; 5 acteurs (auteur, rédacteur, expéditeur, destinataire, créateur) ; 5 types de contextes (juridico-administratif, contexte de provenance, contexte procédural, documentaire, technologique) [The InterPARES 3 Project, TEAM Italy, *Guidelines and Recommendations for E-Mail Records Management and Long-Term Preservation*, 2011, p. 15-17.].

2.2.1. Que collecter ?

Toutes les recherches s'accordent pour **affirmer le rôle essentiel de l'évaluation dans l'élaboration d'une stratégie d'archivage des messages électroniques**. Les solutions techniques, quelles qu'elles soient, ne suffisent pas. Il convient de définir le périmètre des informations à archiver, l'événement déclencheur de l'archivage, la durée de conservation et le sort final des informations à archiver. Ces points doivent être confrontés aux contraintes juridiques existantes¹³⁸.

L'évaluation du périmètre d'archivage se fait à trois niveaux différents : au niveau de l'organisation administrative et fonctionnelle du service ; au niveau du compte de messagerie lui-même ; et enfin au niveau du message.

En revanche, l'adaptation de la stratégie d'archivage aux comportements des utilisateurs est rarement étudiée.

Quelles messageries ?

Le niveau hiérarchique, les fonctions ainsi que les comportements des utilisateurs influent considérablement sur l'intérêt des messageries et des informations qu'elles véhiculent. La littérature professionnelle n'est cependant guère diserte sur ce sujet. Seul le rapport du groupe InterPARES l'évoque, indiquant que toute organisation doit **identifier les comptes de messagerie dont l'archivage est important**, sans pour autant définir des critères de sélection, organisationnels comme fonctionnels¹³⁹. La direction des Archives de France, dans sa note d'information de 2009, propose de cibler de préférence les messageries de certains niveaux hiérarchiques, susceptibles de contenir des informations stratégiques, sous réserve que l'authenticité des informations soit assurée¹⁴⁰. Il peut être cependant utile de **combiner ces recommandations avec une analyse du contexte de production de l'information et des habitudes de travail des décideurs**, qui peuvent être amenés à confier leur(s) messagerie(s) à leur assistant ou leur secrétaire par exemple.

138 BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 23. La FedISA ajoute que toute bonne solution d'archivage doit garantir un accès aisé au contenu, une optimisation du stockage, le respect des exigences en matière de conformité, l'efficacité opérationnelle, une administration efficace, un environnement sécurisé, une conservation pérenne et le suivi des durées de conservation, avec un soin particulier à l'interface utilisateur (*ibid.*, p. 31).

139 The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 25.

140 Note d'information DITN/RES/2009/007 du 3 juin 2009, p. 4.

Quels messages ?

En revanche, la sélection des messages au sein d'un compte de messagerie a fait l'objet d'un grand nombre de recommandations. La terminologie varie certes en fonction des pays : messages à valeur archivistique contre messages sans valeur archivistique¹⁴¹ ; messages à valeur administrative, juridique, financière ou historique ; messages à valeur éphémère et courriels sans valeur¹⁴².

Néanmoins, d'un point de vue général, la littérature professionnelle s'accorde sur le fait que ne doivent pas être archivés :

- ♣ les messages personnels ;
- ♣ les *spams* ;
- ♣ les listes de diffusion ;
- ♣ les messages reçus pour information ;
- ♣ les messages à faible valeur informationnelle ;
- ♣ les messages dits transitoires ;
- ♣ les brouillons¹⁴³.

Seuls les messages, expédiés comme reçus, ayant une valeur informationnelle et/ou juridique pour l'organisation, doivent être conservés¹⁴⁴.

141 REPUBLIQUE ET CANTON DE GENEVE. *Gestion des messages électroniques au sein de l'administration cantonale. Traitement archivistique des messages. Bonnes pratiques concernant l'usage de la messagerie. Logiciel Livelink d'archivage intermédiaire*. Genève, 2013, p. 8-9.

142 BIBLIOTHEQUE ET ARCHIVES NATIONALES DU QUEBEC, Orientations pour la gestion documentaire des courriels du gouvernement du Québec, 2009, p. 6-8.

143 <http://archives.un.org/ARMS/content/managing-emails-records> (consulté le 12/03/2013). BOUDREZ Filip, *Filing and archiving e-mail*. Antwerp, 2006, notamment p. 10. SOYEZ Sébastien, *Directives pour la gestion et l'archivage numérique des e-mails*, Archives générales du Royaume – Algemeen Rijksarchief, version 1.1, octobre 2009, p. 27. FERRANTE Ricardo, SCHMITZ FUHRIG Lynda. *Digital Preservation: Using the Email, Account XML Schema*. Washington, [2007], p. 2-3.

144 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 20.

Le Conseil international des archives, dans son guide pour la gestion des messages électroniques, a élaboré le schéma ci-dessous, afin de déterminer si un message devait intégrer ou non un système d'archivage¹⁴⁵ :

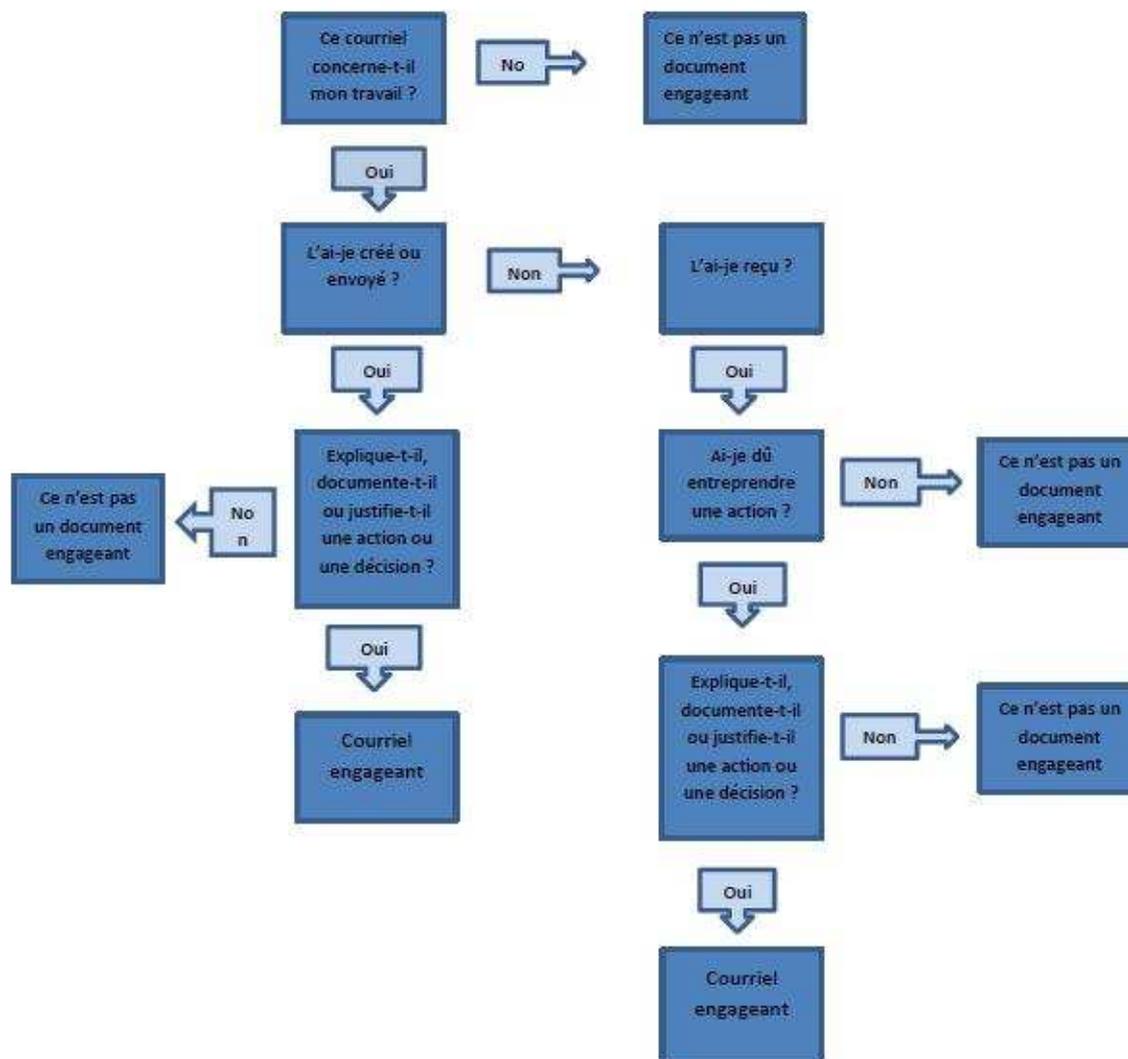


Fig. 4 : schéma d'identification des messages à conserver (Source : Conseil international des archives)

145 CONSEIL INTERNATIONAL DES ARCHIVES (ICA). *L'archivage, clé d'une bonne gouvernance*, module 17 : *Gérer les courriels*, juillet 2012, p. 4

Seuls les utilisateurs finaux sont à même de procéder à cette première sélection, en évaluant la pertinence et l'importance revêtues par chaque message. Logiquement, les Archives de l'État en Belgique comme le groupe InterPARES insistent sur le fait que les organisations doivent accompagner les utilisateurs en édictant des procédures claires, afin que le tri et la conservation des messages soient faits en conformité avec la réglementation en vigueur¹⁴⁶.

Quelles informations ?

La sélection des informations à archiver pour chaque message a également été largement étudiée, notamment au regard des problématiques d'authenticité et d'intégrité. La nature composite des messages électroniques implique de prendre en compte, au moment de l'archivage, l'ensemble de ses composants, afin d'assurer une collecte de qualité¹⁴⁷ :

- ▲ en-tête ;
- ▲ corps ;
- ▲ signature en pièce jointe ;
- ▲ pièces jointes ;
- ▲ fil de discussion ;
- ▲ liens vers des ressources extérieures.

Or la collecte de l'ensemble de ces composants n'est pas sans poser un certain nombre de difficultés soulignées, notamment par le groupe InterPARES, au regard des critères définis par le groupe en matière d'authenticité des documents électroniques :

- ▲ **la fixité du contenu** : un message peut contenir des liens vers des ressources extérieures. La présence de ces liens peut poser de sérieux problèmes lorsque les ressources vers lesquelles pointent ces liens sont essentielles à la compréhension du message et qu'elles ne peuvent être collectées avec le message lui-même. Quant aux pièces jointes, elles peuvent être difficiles à localiser lors d'une opération de migration et peuvent être enregistrées dans des formats obsolètes ;
- ▲ **l'identification de l'action** à laquelle le document participe : un message électronique peut concerner plusieurs actions et peut résulter du mélange de précédents messages. De nombreux classements sont dès lors possibles. La gestion de ces messages dans un système de *records management* peut donc s'avérer problématique ;
- ▲ **les liens avec les autres ensembles documentaires** : un message a une structure qui rend complexe l'établissement du lien entre lui et les autres documents associés à un même processus. Dans certains cas, un unique message peut être associé à plusieurs documents ; dans d'autres cas, un unique document peut nécessiter plusieurs messages pour être expédié en raison de la limitation de la taille des pièces jointes. Reconstituer une chaîne de messages peut par ailleurs s'avérer complexe en raison des protocoles de filage utilisés par les logiciels retenus ;
- ▲ **l'identification des acteurs** : ils peuvent parfois être difficile voire impossible à identifier dans le cas d'un message électronique, en raison des choix organisationnels et techniques effectués par une organisation. Des procédures doivent être mises en place pour identifier

146 SOYEZ S., *Directives pour la gestion et l'archivage numérique des e-mails*, Archives générales du Royaume – Algemeen Rijksarchief, version 1.1, octobre 2009, p. 29. The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 26.

147 *Ibid.*, p. 6 et 49. SOYEZ S., *Directives pour la gestion et l'archivage numérique des e-mails*, Archives générales du Royaume – Algemeen Rijksarchief, version 1.1, octobre 2009, p. 31-32. PROM C. J., *Op. cit.*, p. 10-11.

correctement les personnes physiques impliquées dans l'envoi d'un message ;

- ♣ **L'identification du contexte de création et de gestion** : les messages ne fournissent que des données relatives à leur contexte technologique. Les choix opérés pour l'installation et la gestion des serveurs de messagerie (format des en-têtes, structure du stockage), rarement documentés, peuvent en effet affecter la conservation et la valeur des messages. Seules des actions ciblées permettent de garantir la disponibilité des informations nécessaires pour décrire de manière adéquate les messages¹⁴⁸.

Le projet InSPECT a défini trois scénarii permettant de qualifier le degré d'authenticité d'un message électronique :

- ♣ le premier, dit « **scénario noyau** », indique les données minimales à collecter pour établir l'authenticité et l'intégrité d'un message électronique. Avec ces données, il est possible d'établir l'origine d'un message, les étapes de sa transmission (pour les messages reçus), ses destinataires, son objet, son corps et l'existence des pièces jointes ;
- ♣ le deuxième scénario, dit « **scénario fil de discussion** » permet de replacer un message dans un fil de discussion. Il permet de replacer un message dans un contexte de discussion ;
- ♣ le troisième, dit « **scénario idéal** », recense toutes les données à collecter.

Le tableau ci-dessous synthétise les différents scénarios¹⁴⁹ :

Champ	Description	Scénario noyau	Scénario fil de conversation	Scénario idéal
Local-part	Adresse	X	X	X
Domain-part	Nom de domaine	X	X	X
Domain-literal	Adresse IP			X
Relationship	Relation de l'acteur avec le message : auteur, expéditeur, destinataire pour action, pour information ou pour information caché	X	X	X
Subject	Objet	X	X	X
Trace-field	Trace de la réception	X	X	X
Attachments	Pièces jointes	X	X	X
Message-ID	Identifiant unique		X	X
References	Ensemble des messages parents dans le fil du message		X	X
Sent-Date	Date d'expédition			X
Received date	Date de réception			X
Display Name	Nom affichable des adresses de messagerie			X
In-reply-to	Message parent immédiat			X
Keywords	Mots-clés			X
Message body	Corps	X (sans mise en page)	X (sans mise en page)	X (avec mise en page)

148 The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 15-17.

149 KNIGHT J. *InSPECT Project Document - Significant Properties Testing Report : Electronic Mail*, Londres : JISC/The National Archives/King's College London, 2009, p. 15-16.

Le degré d'intégrité et d'authenticité souhaité joue donc un rôle non négligeable dans la définition de la stratégie et des opérations d'archivage. Le groupe InterPARES recommande de porter une attention particulière aux points suivants :

- ⤴ **l'unicité** : tout *record* envoyé par messagerie électronique doit ainsi être intégré dans le système d'archivage comme une entité distincte. Tout message doit également être considéré comme un *record* ;
- ⤴ **l'identification** : tout message à valeur de *record* doit disposer d'un identifiant unique dans le système d'archivage et doit être classé et associé à une durée de conservation et à un sort final ;
- ⤴ **l'identification des acteurs** : l'adresse de l'expéditeur doit correspondre à une personne identifiable. À défaut, l'expéditeur sera réputé inconnu. Dans le cas où le lien entre un individu et une adresse est fait par un outil externe (base de données, tableur), celui-ci doit être maintenu et pris en compte lors de l'élaboration de la stratégie d'archivage. Une attention particulière devra être portée aux listes de diffusion pour permettre d'identifier les personnes incluses dans celles-ci ;
- ⤴ **le contexte** : toutes les informations de l'en-tête doivent être reprises ;
- ⤴ **la datation** : les dates d'expédition, de réception, et éventuellement l'heure doivent être intégrés dans le profil des métadonnées¹⁵⁰.

2.2.2. Comment et quand collecter ?

Deux méthodes de collecte peuvent être identifiées :

- ⤴ **la collecte globale**, qui s'intéresse à un écosystème de messageries ;
- ⤴ **la collecte ciblée**, qui s'intéresse aux comptes ou aux messages individuels¹⁵¹.

La collecte globale

Certaines entreprises, notamment dans le secteur de la banque, de l'assurance ou de la santé, prenant conscience des risques légaux, ont fait le choix de gérer leurs messages électroniques en dehors de l'infrastructure de serveurs de messageries en utilisant des logiciels, propriétaires comme open source (ex. : Symantec Entreprise Vault, Iron Mountain Nearpoint, Smarsh Email Archiving Suite, Mail Archiva), assurant une capture automatique des messages produits et reçus au point de transmission, ainsi que leur transfert dans un espace de stockage externe où les utilisateurs ne peuvent les détruire. Les utilisateurs accèdent aux messages *via* un système de pointeurs. Le stockage des messages peut être hiérarchisé pour garantir une performance et une disponibilité adaptées aux besoins de restitution des utilisateurs, permettant ainsi d'optimiser les coûts de sauvegarde et de restauration. Ce type de système permet de paramétrer des durées de conservation fondées sur des critères définis (expéditeur, destinataire, objet, date, mots-clés) et de détruire

150 The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 35-43 et SOYEZ S., *Directives pour la gestion et l'archivage numérique des e-mail (...)*, p. 30.

151 PROM, Chris, « Email preservation options », *Practical E-records blog*, 17 novembre 2011, disponible en ligne : <http://e-records.chrisprom.com/email-preservation-options> (consulté le 27 mars 2013). DIRECTION INTERMINISTERIELLE DES SYSTEMES D'INFORMATION ET DE COMMUNICATION (DISIC). *Archivage électronique. Un nouveau domaine d'expertise au service de la gouvernance des systèmes d'information*, DISIC, 2012, t. 1, p. 14. Disponible en ligne : http://references.modernisation.gouv.fr/sites/default/files/DISIC_AE_-_Guide_bonnes_pratiques_0.PDF (consulté le 8 mars 2013).

systématiquement les messages dont la durée de conservation est dépassée.

Cette approche globale permet aux organisations qui l'ont retenue de limiter les risques liés à une trop grande responsabilité donnée dans la conservation de documents engageants aux utilisateurs finaux, ainsi que d'éviter les coûts induits par la mise en place de systèmes de *records management*¹⁵². Elle présente néanmoins d'importants inconvénients : impossibilité de gérer finement les droits et de cibler les messages collectés ; risque de non-conformité avec la réglementation sur le respect de la vie privée (voir ci-dessus). À titre d'exemple, la FedISA ne se montre pas pleinement convaincue de son opportunité¹⁵³. Elle peut cependant être utilisée dans le contexte français pour l'archivage des messageries officielles, où les tris seront faibles ou inexistantes.

La collecte ciblée

Le stockage de messages dans un répertoire au sein du client de messagerie ne constitue qu'une solution temporaire, qui pose les problèmes suivants :

- ▲ durabilité et pérennité, notamment compte tenu du format d'export propriétaire ;
- ▲ perte d'informations lors des opérations « d'archivage » proposées par certains de ces logiciels ; accès limité au seul utilisateur ;
- ▲ stockage limité ;
- ▲ caractère corrompible, instable et peu fiable des dossiers stockés localement.

La littérature et la pratique recommandent donc l'export des messages hors du client de messagerie, en s'appuyant sur les procédures du *records management*¹⁵⁴.

Collecte au fil de l'eau (message par message)

Plusieurs services d'archives se sont efforcés de mettre au point des outils permettant de gérer les messages conformément aux exigences du *records management* et d'isoler les messages qui doivent être archivés par l'entité. Cette politique assure une collecte au fil de l'eau des messages engageant la responsabilité de l'entité.¹⁵⁵

On peut citer à titre d'exemples les expériences menées par la ville d'Anvers, par les Pays-Bas et par le canton de Genève :

- ▲ le projet DAVID (ville d'Anvers) : les messages sont rendus exportables depuis le client de messagerie vers un serveur de fichiers, après enregistrement des métadonnées de transmission et de contexte¹⁵⁶ ;
- ▲ le projet expérimental des Pays-Bas : un module complémentaire à Outlook permet aux utilisateurs, *via* une interface adaptée, de saisir un nombre limité de métadonnées complémentaires nécessaires à la contextualisation des messages officiels (dossier, fonction). Les messages sont générés au format XML nativement et sont expédiés avec une feuille de style qui permet de convertir les fichiers au format html au moment de la

152 PROM C. J., *Preserving Email (...)*, p. 17 et 25.

153 BALLETT P., RIETSCH J.-M., *Op. cit.*, p. 11 et 30.

154 SOYEZ S., *Directives pour la gestion et l'archivage numérique des e-mails (...)*, p. 31.

155 Note d'information DITN/RES/2009/007 du 3 juin 2009, p. 4 et The InterPARES 3 Project, TEAM Italy, *Guidelines and Recommendations for E-Mail Records Management and Long-Term Preservation*, 2011, p. 12, 23.

156 BOUDREZ Filip, *Filing and archiving e-mail (...)*, p. 2-9.

- réception. Ils peuvent ensuite être récupérés pour être archivés¹⁵⁷ ;
- ♣ le projet Livelink (République et canton de Genève) : un dossier Livelink est ajouté à l'arborescence de la boîte aux lettres des utilisateurs du client Outlook. Il donne accès à deux espaces dédiés aux archives : un espace « personnel-professionnel » mis à disposition de chaque utilisateur, qui est libre d'y créer sa propre arborescence ; un espace « entreprise », partagé entre plusieurs utilisateurs et structuré selon le cadre de classement et le calendrier de conservation du service élaborés avec l'archiviste et les Archives d'État. Ce second espace n'est déployé que sur demande expresse du service¹⁵⁸. L'outil, accessible depuis la messagerie et depuis un portail web, permet d'archiver les messages par simple « glisser-déposer » dans une infrastructure centralisée disposant d'une fonction de recherche. Le stockage dans cet espace est illimité. Les messages enregistrés dans l'espace entreprise peuvent ensuite être récupérés par le service d'archives.

Cette approche place **l'utilisateur, qu'il soit l'expéditeur ou le destinataire du message, au centre du dispositif d'archivage**, puisque c'est lui qui décide quels messages ont une valeur pour l'entité. Elle a notamment l'avantage de garantir le respect de sa vie privée. La question principale consiste à définir les justes contraintes à faire peser sur l'utilisateur, afin qu'il devienne acteur non contraint mais volontaire du système. Aussi est-il indispensable de ne bouleverser qu'*a minima* ses habitudes de travail, en automatisant au maximum le processus. De son implication et de sa rigueur dépendent néanmoins la réussite des opérations d'archivage et la capture des documents engageants pour l'entité. La conduite du changement et l'accompagnement des utilisateurs (création de modèles de messages avec les métadonnées pré-remplies, mise en place de fonctionnalités d'alerte, utilisation des ressources de l'analyse sémantique), lors de l'installation d'un environnement logiciel, sont donc primordiales¹⁵⁹.

D'un point de vue technique, la littérature recommande d'interfacer tout ou partie des comptes de messagerie à des *Electronic Records Management Systems* (ERMS). La collecte et le classement dans l'ERMS doivent être au maximum automatisés – par exemple sur la base de modèles de documents, du contexte ou de l'analyse sémantique du contenu - et préserver l'intégrité et l'authenticité du message. La solution doit laisser à des utilisateurs la possibilité d'enrichir ou de corriger les métadonnées. Cette solution impose aux organisations de planifier les ressources pour acquérir, configurer et exploiter ces ERMS, en plus des politiques et procédures adaptées. De tels systèmes sont notamment recommandés pour les industries et des agences centralisées et sensibles aux questions de contentieux¹⁶⁰. Néanmoins, Christopher Prom estime que la plupart des tentatives de collecte au fil de l'eau reposant sur le bon vouloir des utilisateurs se sont révélées difficiles à mettre en oeuvre. Reprenant les conclusions d'autres études et présentations, il indique également que les tentatives de capture dans des ERMS se sont révélées peu efficaces, faute d'adhésion des utilisateurs aux outils implémentés¹⁶¹.

157 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 48-51.

158 REPUBLIQUE ET CANTON DE GENEVE. *Gestion des messages électroniques au sein de l'administration cantonale. Traitement archivistique des messages. Bonnes pratiques concernant l'usage de la messagerie. Logiciel Livelink d'archivage intermédiaire*. Genève, 2013, p. 20.

159 CLUB DES RESPONSABLES DE POLITIQUE ET PROJETS D'ARCHIVAGE [CR2PA], *Op. cit.* p. 11 et The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p.63.

160 The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 25-26. CLUB DES RESPONSABLES DE POLITIQUE ET PROJETS D'ARCHIVAGE [CR2PA], *Op. cit.*, p. 13-14. SOYEZ S., *Directives Op. cit.*, p. 35-36. The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 30-33. PROM C. J., *Preserving Email (...)*, p. 26. BALLEP P., RIETSCH J.-M., *Op. cit.*, p. 30-31.

161 PROM C. J., *Preserving Email (...)*, p. 15.

Collecte *a posteriori* (compte par compte)

Faute de disposer des moyens, notamment logiciels, de mettre en place une collecte ciblée, de nombreuses organisations et des services d'archives ont dû envisager la collecte *a posteriori* de comptes de messageries. Il s'agit dans ce cas de récupérer à des échéances plus ou moins régulières, notamment lors du départ d'une personne (mutation par exemple), l'ensemble des messages qu'elle a produits et reçus dans le cadre de ses activités avec la messagerie mise à disposition par son employeur.

En France, une collecte *a posteriori* a, par exemple, été réalisée par plusieurs missions des Archives de France placées auprès des ministères (Mission auprès des ministères sociaux, Mission auprès du ministère de la Culture et de la Communication) à l'occasion des opérations d'archivage menées dans les cabinets ministériels lors du changement de majorité intervenu à la suite des élections présidentielles du printemps de l'année 2012. Aux États-Unis, le NARA vient d'adopter une politique de collecte fondée sur cette approche, comme en témoigne le projet¹⁶².

Si cette solution a le mérite de la simplicité, elle présente néanmoins plusieurs inconvénients en raison de l'absence de tri effectué dans la messagerie collectée, ce qui entraîne la présence de messages relatifs à la vie privée de la personne concernée et de nombreux messages n'ayant aucun intérêt à être archivés : brouillons, listes de diffusion, *spams*, messages reçus pour information. Seules des messageries identifiées comme particulièrement critiques pour l'organisation en raison de la qualité ou des méthodes de travail des personnes concernées peuvent faire l'objet d'une telle collecte, étant donné l'importance des opérations de tri et de classement que leur collecte induit. Cette problématique du traitement de messageries collectées a d'ailleurs fait l'objet d'une première analyse par le projet CERP¹⁶³.

2.2.3. Comment traiter et conserver ?

Traitement intellectuel

Le traitement intellectuel et la description sont certainement les thématiques les moins nettement interrogées par la littérature professionnelle. L'adaptation de la pratique archivistique aux messageries électroniques pose cependant plusieurs problèmes.

Les messageries officielles peuvent être assimilées à des minutiers chronologiques et peuvent être traitées comme tels dans des instruments de recherche classiques respectant la norme ISAD (G). Dans ce cas :

- ▲ l'émetteur de la messagerie est considéré comme le producteur (niveau fonds ou sous-fonds) ;

162 NATIONAL RECORDS AND ARCHIVES ADMINISTRATION. *Bulletin 2013-02 "Guidance on a New Approach to Managing Email Records" draft*. Disponible en ligne : <http://www.archives.gov/records-mgmt/bulletins/2013/2013-02.html> (consulté le 3 septembre 2013).

163 FERRANTE Ricardo, SCHMITZ FUHRIG Lynda. *Digital Preservation: Using the Email, Account XML Schema*. Washington, [2007], p. 2-3.

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

- ⤴ la messagerie est considérée comme une série organique ;
- ⤴ chaque message est considéré comme une pièce, avec ses annexes (pièces jointes).

Des dossiers intermédiaires peuvent être constitués en rassemblant des messages selon un critère chronologique (année/mois/semaine/jour).

Un traitement similaire des messageries interpersonnelles est moins facile à concevoir pour plusieurs raisons :

- ⤴ si les utilisateurs disposent de plusieurs messageries, il y aurait autant de minutiers chronologiques que de messageries. Serait-il pour autant pertinent de définir autant de sous-séries que de comptes de messageries ? ;
- ⤴ dans bien des cas, des systèmes de classement existent et une conservation des messages « à plat », sous forme de minutiers chronologiques, entraînerait un appauvrissement des messages collectés, faute d'exploitation de la richesse des métadonnées collectées.

Quelles sont les autres solutions disponibles ?

Une première solution pourrait consister, avec l'aide d'un client de messagerie, à réaliser un **classement manuel** des messages collectés dans les différents comptes avec un plan de classement unique sous forme de dossiers locaux. Il serait alors possible de disposer d'une arborescence unique à base de dossiers et de sous-dossiers conçus selon une logique adaptée au fonds d'archives (thématique, géographique, nominative). Le recours à une telle solution permettrait alors d'éliminer les messages sans intérêt, les doublons créés par le simple transfert de confort entre comptes de messagerie et de produire une description compatible avec la norme ISAD (G) : les différents niveaux de l'arborescence constitueraient des dossiers et des sous-dossiers, tandis que chaque message serait considéré comme une pièce.

Cependant, cette solution présente plusieurs inconvénients :

- ⤴ Techniquement chronophage, elle ne peut être envisagée à l'échelle globale d'une organisation pour le traitement des messageries de plusieurs centaines de personnes. Le projet CERP a ainsi évalué à 10 heures le traitement manuel de 6 000 messages et a estimé le classement de boîtes non organisées irrationnel. Les institutions partenaires ont donc décidé de conserver la structure originelle des boîtes collectées, sans reclassement aucun¹⁶⁴ ;
- ⤴ Juridiquement, elle risque de porter atteinte à l'intégrité du fonds collecté, surtout si le producteur a défini des exigences particulièrement avancées en la matière. Toute modification du classement reviendrait alors à modifier les métadonnées de l'en-tête des messages et à porter atteinte à l'intégrité du compte ;
- ⤴ Informationnellement parlant, elle ne permet pas de tirer parti de la richesse des métadonnées existantes dans les en-têtes, qui ne peuvent toutes être récupérées dans des balises EAD.

Si l'élaboration d'un instrument de recherche traditionnel conforme à la norme ISAD (G) est tout à fait envisageable¹⁶⁵, il faut cependant bien être conscient qu'il ne constitue qu'un mode d'accès parmi d'autres aux messages électroniques. Toute stratégie d'accès à des messages électroniques collectés devra être capable de tirer parti des métadonnées présentes dans leurs en-têtes, voire d'utiliser des outils d'indexation automatique.

164 *Ibid.*, p. 5. ADGENT Nancy. *Collaborative Electronic Records Project : An Introduction and Overview*. Sleepy Hollow, New York et Washington, 2008, p. 11.

165 http://siarchives.si.edu/cerp/CERP%20SIA%20EAD_CC.PDF. FERRANTE R., SCHMITZ FUHRIG L., *Op. cit.*, p. 3.

Traitement technique

Plusieurs solutions ont été étudiées pour assurer la pérennisation des messages électroniques en garantissant l'intégrité, l'authenticité, la fiabilité et l'accessibilité des messages : l'impression, l'émulation, la migration. Seule la dernière est cependant conseillée pour la conservation à long terme.

La simple impression, qu'elle soit papier ou au format PDF, ne constitue pas en-soi un archivage adéquat, car elle restreint le nombre d'informations (notamment les métadonnées contextuelles) qui seront conservées. Elle n'est conseillée qu'à titre très provisoire, à défaut d'archivage électronique¹⁶⁶. L'impression en PDF n'est pas plus recommandée, pour les mêmes raisons de pertes d'informations dans l'en-tête. La seule information conservée est une image fixe du message tel que représentée par le client¹⁶⁷.

L'émulation a également été rejetée par les différentes études, au départ pour un manque de maturité¹⁶⁸, puis en raison de la trop grande diversité des composants des messages¹⁶⁹.

La migration est donc recommandée. Les recherches menées par les différentes équipes qui se sont intéressées à la conservation des messages ont toutes conclu à l'intérêt que pouvait représenter le format XML, en raison de son ouverture, de son caractère lisible par un homme comme par une machine, de la souplesse de structuration des données et métadonnées qu'il permet et de son indépendance vis-à-vis de toute plate-forme matérielle ou logicielle. Son utilisation permet également de mettre en place des stratégies de conservation différenciées pour les différents composants des messages, tout en préservant le lien entre ceux-ci¹⁷⁰.

Plusieurs institutions ont donc défini des schémas XML.

Projet expérimental néerlandais

Ce projet a conçu un cadre XML pour la conservation des messages électroniques, au niveau unitaire. Le message transformé a une structure plus complexe que le fichier originel, mais est également plus riche.

L'AIP constitué à partir du message comprend quatre éléments liés les uns aux autres :

- ▲ le **fichier XML (a)** : composant principal, il est produit soit par conversion du fichier de transmission (*transmission file*), soit immédiatement avec le client de messagerie pour les messages expédiés. Il contient des informations provenant de plusieurs sources :
 - métadonnées de l'en-tête (expéditeur, objet, date, composition du message) ;
 - métadonnées ajoutées par l'utilisateur (dossier et fonction) ;
 - date de conversion.

166 SOYEZ S., *op. cit.*, p. 31.

167 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 33-34. FERRANTE R., SCHMITZ FUHRIG L., *Op. cit.*, p. 6.

168 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 40-46.

169 The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 54.

170 SOYEZ S., *Op. cit.*, p. 21-22 et 36-39. The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 48-49. PROM C. J., *Preserving Email (...)*, p. 23-24.

Toute organisation peut rajouter des métadonnées.

Le corps (b) et les pièces jointes (c) peuvent être soit disjoints en tant que fichiers séparés vers lesquels pointe le fichier XML, soit intégrés directement dans le fichier, avec une référence à la feuille de style utilisée pour réaliser la version html de réception – c'est le cas pour les messages expédiés au moyen du module complémentaire à Outlook. Lorsque le corps est stocké dans un fichier séparé, le fichier XML doit préciser les informations suivantes : nom du fichier, type du fichier, localisation du fichier. Pour les messages expédiés avec le module complémentaire à Outlook, une copie des pièces jointes peut être ajoutée au fichier XML stocké. Dans les autres cas, les pièces jointes doivent être stockées dans leur format natif, avec une référence dans le fichier XML et les informations suivantes : nom de fichier ; type de fichier ; localisation du fichier ;

- ♣ **les métadonnées complémentaires (d)** : la conservation à long terme nécessite un fichier de métadonnées extérieur dont le contenu précis est décidé par chaque institution. Il a notamment pour but de stocker toutes les informations de l'en-tête non reprises dans le fichier XML ;
- ♣ **le journal de la conservation (e)** : il contient toutes les informations sur les opérations de conservation réalisées sur le message. Il peut également contenir des informations sur les exigences en matière d'accès et de conservation. Ce fichier a deux objectifs : la vérification de l'authenticité et de l'accessibilité ; la gestion de chaque composant du message, sans mettre en danger l'authenticité du message lui-même en tant que document. Ce journal est initié lors de la conversion initiale en XML. Il doit être conçu pour être aisément et régulièrement mis à jour. Il peut s'agir soit d'une base de données, soit d'un fichier XML. Sont successivement précisés : la forme initiale du document lors de la conversion ; toutes les informations utiles sur la conversion, y compris le nom du logiciel utilisé, la date et le moment de la conversion et le format du nouveau fichier. Il doit être mis à jour à chaque action de préservation ;
- ♣ **le fichier de transmission originel (f)** : il s'agit de la source la plus complète et la plus fiable. Si le processus de conversion échoue, il permet de faire des vérifications, notamment dans l'en-tête. Il peut également servir de base à de nouvelles stratégies de conservation.

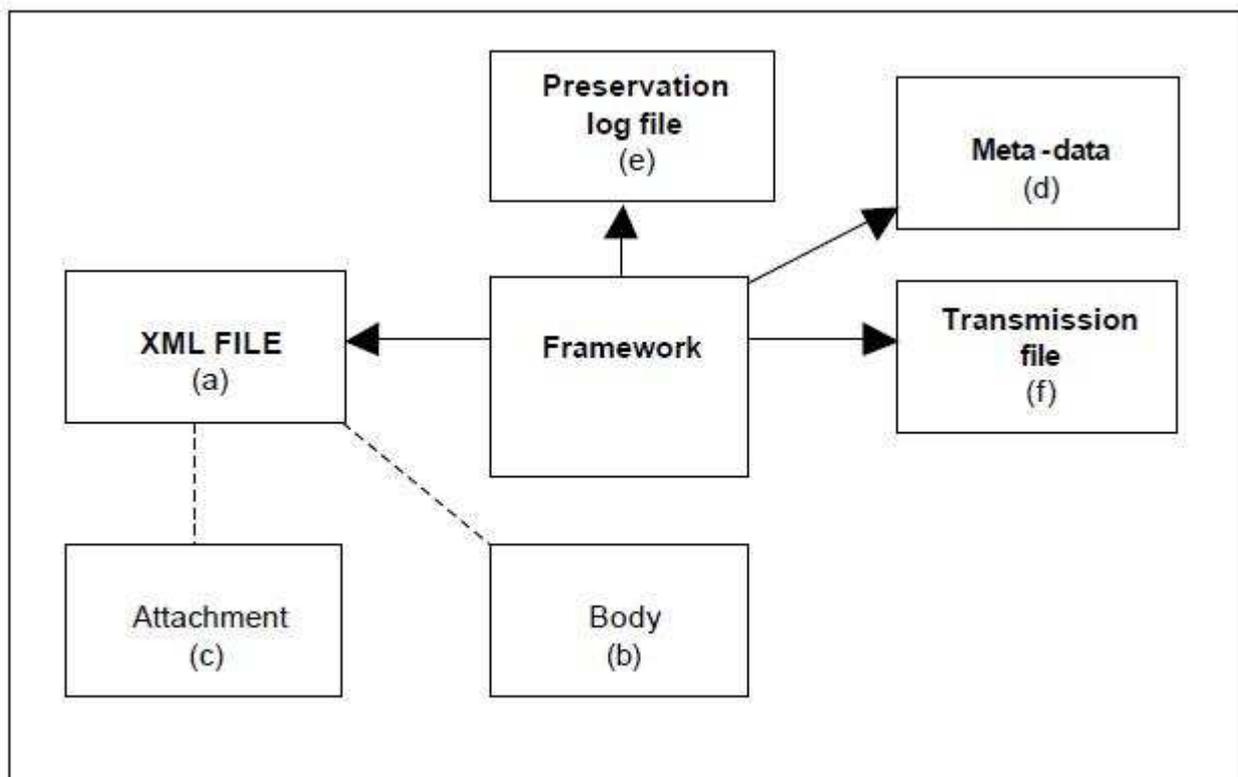


Fig. 5 : modélisation de l'AIP correspondant à un message électronique (Source : Projet Testbed)

Les métadonnées du fichier XML peuvent aisément être traitées et intégrées dans un fichier ead ou dans le système de gestion des métadonnées propre au service d'archives concerné¹⁷¹.

Projet CERP

Le projet CERP a, lui, réfléchi, non plus au niveau du message unitaire mais au niveau du compte de messagerie lui-même.

Les comptes de messagerie, une fois collectés avec un minimum de métadonnées sous forme d'un bordereau de versement, sont documentés, contrôlés au moyen d'un anti-virus et convertis au format mbox. Les pièces jointes sont extraites au moyen du logiciel EZDetach et analysées au moyen de JHOVE et de DROID. Leur traitement n'est pas sans poser quelques difficultés : fixation des dates automatiques, gestion des doublons, gestion des données personnelles¹⁷². Un fichier XML est généré à l'aide d'un outil développé spécifiquement, le CERP Parser. En parallèle, un fichier METS est créé. L'AIP constitué contient :

- ♣ le compte ou le message dans son format d'origine (msg, pst) ;
- ♣ le compte au format XML (format de conservation) ;
- ♣ un fichier zip conservant le compte de messagerie au format mbox, les messages d'erreur, un résumé et les pièces jointes ;
- ♣ les métadonnées archivistiques (sort final), descriptives (instrument de recherche) et techniques (rapport d'analyse des pièces jointes) ;

171 DIGITAL PRESERVATION TESTBED. *Op. cit.*, p. 53-58

172 FERRANTE R., SCHMITZ FUHRIG L., *Op. cit.* p. 4-5.

- ▲ les logs du parser au format .zip ;
- ▲ l'instrument de recherche ;
- ▲ le fichier METS utilisé pour l'entrée dans le dépôt de stockage ;
- ▲ les métadonnées administratives et descriptives encodées au format METS ;
- ▲ une feuille de style au format XML pour la communication¹⁷³.

AIP Assembly

```
Source e-mail (.pst, .msg, etc.)
Preservation e-mail
  Parsed messages (XML)
Parser Directory Tree.zip
  Pre-parsed messages (mbox)
  Bad messages
  Message Summary
  Attachments
Subject-sender log.zip (from parser output)
Metadata narrative.zip – complete (administrative & descriptive)
Finding aid narrative.zip
METS.xml
File Name METS.xml (administrative & descriptive metadata encoded in METS)
Display aids
  XML stylesheet
```

Fig. 6 : Schéma de l'AIP correspondant à un compte de messagerie (Source : CERP, *Email preservation workflow*, 2010, p. 2.)

L'E-Mail Account Schema est une représentation complète d'un compte de messagerie, jusqu'au niveau le plus fin du message. Le schéma reprend sous forme de balises l'organisation en dossiers et sous-dossiers. La mise en regard du schéma avec la RFC 2822 a montré que toutes les informations étaient bien préservées. Les pièces jointes sont stockées en dehors du fichier XML. La granularité facilite l'accessibilité et la compréhension en permettant des stratégies avancées de recherche, notamment grâce à la préservation intégrale des en-têtes¹⁷⁴.

2.2.4. Comment restituer ?

L'interrogation et la restitution de messageries électroniques doivent être possibles tant au niveau unitaire qu'au niveau global¹⁷⁵. Elle **doit porter tout autant sur les métadonnées de l'en-tête que sur le corps du texte**, fonctionnalité assurée de manière standard par les clients de

173 ADGENT N., *Op. cit.*, p. 12-14. FERRANTE R., SCHMITZ FUHRIG L., *Op. cit.*, p. 2.

174 *Ibid.*, p. 6-7.

175 *Ibid.*, p. 2.

messagerie. Dans le cas où la collecte s'effectue au niveau d'un outil de collecte globale ou d'un ERMS, la restitution est souvent permise soit directement *via* le client de messagerie, soit *via* l'interface de recherche de l'outil lui-même.

La communication des messages conservés au format XML n'est pas sans poser de problème. Deux solutions principales peuvent être proposées :

- ▲ le **stockage d'une version des messages dans un des formats standards**, par exemple le format eml, ce qui permet l'utilisation, pour leur restitution, d'un client de messagerie traditionnel. Cette solution pose néanmoins plusieurs problèmes : viabilité à long terme ; stockage en multiples exemplaires d'un même message ;
- ▲ le **stockage d'une feuille de style dans l'AIP**, ce qui permet une restitution sous forme graphique des comptes de messagerie archivés. C'est la solution retenue dans le cadre du projet CERP (cf. ci-dessus).

Selon Christopher Prom, le développement d'outils de navigation, d'interrogation et de restitution de messages conservés au format XML constitue un chantier à ouvrir¹⁷⁶.

Conclusion

Si les règles de bonne gestion des messageries ont fait l'objet de nombreux guides, la question du traitement archivistique est moins bien cernée, notamment pour tout ce qui a trait au tri, à la description et au classement des messages. La question de la restitution des messages archivés n'est malheureusement que très peu abordée. Quant à l'approche diplomatique, elle est suffisamment avancée pour permettre une évaluation de l'intégrité des messageries par l'analyse des messages exportées (à partir des *significant properties* par exemple).

Le problème majeur reste celui de la collecte des messageries. Plusieurs stratégies ont été définies par les organisations et les services d'archives, en fonction de leur contexte et de leurs besoins en matière de conservation des documents. Il apparaît cependant que les stratégies reposant sur les utilisateurs ne sont pas sans poser de difficultés et que le sujet redevient d'actualité comme le démontrent les initiatives récentes du NARA (programme Capstone) et des Archives de Nouvelles-Galles du Sud.

176 PROM C. J., *Preserving Email (...)*, p. 37.

III - Les expérimentations menées dans le cadre du volet 2 du POC VITAM

À partir des différentes options de traitement et d'archivage listées dans les deux premières parties de cette étude, l'équipe VITAM a choisi de planifier deux types de tests :

- des tests techniques pour s'assurer de la rigueur de conception des outils listés ;
- des tests métier pour évaluer leur pertinence dans un contexte réel de collecte.

NB : Ces tests ont été réalisés conjointement par le ministère de la Défense (Service historique de la Défense-SHD) et par le ministère de la Culture et de la Communication.

3.1. Présentation générale des tests réalisés

3.1.1. Objectifs généraux

Les tests réalisés dans le cadre du présent projet avaient un triple objectif :

- ♣ vérifier si les constats faits lors de la revue de littérature, parfois anciens, étaient toujours d'actualité et compatibles avec l'environnement technique et logiciel de l'administration française, en se focalisant sur trois points :
 - ♣ les formats recommandés pour l'export ;
 - ♣ les outils logiciels fournis par le marché ou développés dans le cadre des projets de coopération entre institutions de conservation ;
 - ♣ la capacité des exports à restituer les données essentielles à collecter ;
- ♣ commencer à définir, à partir des hypothèses de travail formulées par l'équipe-projet en matière de modèle conceptuel à la date de la rédaction de la présente étude, une stratégie de collecte, de conservation et de restitution des messages et messageries électroniques dans la solution VITAM prenant en compte les particularités des différents ministères :
 - ♣ exports recommandés ;
 - ♣ constitution des paquets à verser dans la solution (*Submission Information Package* ou SIP au sens de la norme OAIS) et identification des traitements préalables à leur constitution ;
 - ♣ constitution des paquets à conserver dans la solution (AIP au sens de la norme OAIS) ;
 - ♣ modalités de restitution des objets archivés depuis la solution VITAM (*Dissemination Information Package* ou DIP au sens de la norme OAIS) ;
- ♣ réfléchir à une structuration des métadonnées afin d'alimenter en jeux de tests la preuve de concept relative à la pertinence de l'utilisation des technologies NoSQL dans la conception de la solution VITAM.

3.1.2. Choix et sélection des boîtes mails

Pour ses tests des jeux, l'équipe-projet a décidé d'utiliser des données exportés depuis des clients de messageries installés sur des postes de travail individuels, similaires à ceux obtenus lors des collectes ciblées *a posteriori* décrites ci-dessus dans la section 2.2.2. Il aurait sans doute été intéressant de comparer les exports obtenus avec ceux obtenus directement à partir de serveurs d'échange. Ce travail n'était cependant pas réalisable dans le délai imparti pour la réalisation de la présente étude.

L'attention des membres de l'équipe-projet s'est naturellement tournée vers les **exports réalisés à partir des clients de messagerie les plus utilisés dans le contexte de l'administration publique française**. Deux clients ont été retenus : Thunderbird, recommandé par la direction interministérielle des Systèmes d'information et de communication (DISIC), et Outlook, client le plus utilisé dans le monde de l'entreprise comme dans celui de l'administration (cf. *supra*). Un sondage effectué par les Archives nationales auprès des missions des Archives de France a confirmé la pertinence de ce choix, même s'il a révélé l'utilisation ponctuelle, mais dans certains cas croissante, d'autres clients : Lotus Note ou OS Mail. Ont été écartés les exports de clients du marché en *webmail*, type Yahoo Mail ou Google Mail. Enfin, des tests complémentaires sur des exports depuis le client de messagerie Lotus Note sont en cours de réalisation par le Service historique de la Défense et seront prochainement disponibles.

Le Service historique de la Défense (SHD) a donc procédé à des exports depuis les comptes de messageries de deux de ses agents, le premier depuis un client Outlook au format pst, le second depuis un client Thunderbird au format mbox d'export du client.

Les messageries officielles utilisées par le ministère de la Défense pour la diffusion des ordres de commandement ont également retenu l'attention de l'équipe-projet, et notamment l'application Melind@, dont le décommissionnement est prévu à brève échéance. Basée sur l'utilisation du client de messagerie Thunderbird, cette messagerie d'autorité présente la particularité de disposer également des fonctionnalités de validation des projets d'ordres, avec plusieurs étapes de relecture, de validation et de signature. Il a semblé utile à l'équipe-projet d'étudier l'impact de la présence de ces fonctionnalités sur les exports effectués à partir du client Thunderbird. Les exports ont été réalisés par l'Etat-major de la Marine à la demande du Service historique de la Défense au format eml. L'analyse des fichiers exportés a néanmoins rapidement montré que les informations relatives au processus d'élaboration, de validation et de signature des ordres étaient intégrées dans le corps même des messages et n'avaient aucune conséquence sur la qualité du fichier eml. Les tests réalisés ensuite sur ces fichiers ont donc été similaires à ceux effectués à partir d'exports de messageries interpersonnelles.

3.1.3. Choix et installation des outils

La phase de revue de littérature a permis à l'équipe-projet d'**identifier plusieurs outils, soit disponibles sur le marché, soit développés par les services patrimoniaux dans le cadre de projets de collaborations**. Les outils recensés par la *Digital Preservation Coalition*, par l'étude InSPECT ainsi que par la *Wellcome Trust Foundation* ont servi de base à l'établissement de la liste

des outils à tester.

Deux catégories d'outils ont été exclues :

- ⤴ les **outils de collecte globale des messages** directement sur les serveurs d'échange, comme Symantec Enterprise Vault, Iron Mountain Nearpoint, Smarsh Email Archiving Suite, Mail Archiva (cf. *supra*). Leur nature propriétaire, les contraintes juridiques liées à leur utilisation ainsi que la complexité de leur installation ont conduit à leur exclusion du champ de l'étude ;
- ⤴ les **modules de collecte ciblée**, greffés à un client de messageries et servant d'interface entre ce client et un ERMS, comme l'outil LiveLink utilisé par les Archives de la République et canton de Genève. Étant donné le faible nombre d'ERMS utilisés dans l'administration française à ce jour, l'étude de ces outils a semblé peu pertinente aux membres de l'équipe-projet.

Ont donc été retenus trois types d'outils :

- ⤴ **deux progiciels du marché**, Aid4Mail et EmailChemmy, étudiés dans le cadre de l'étude de la *Digital Preservation Coalition*, qui assurent la conversion de messages et de messageries dans des formats standard (mbox, eml, .PDF, csv) ;
- ⤴ **CERP Email Parser et Xena**, logiciels développés par des services d'archives, et qui assurent un traitement des messageries permettant de générer une enveloppe XML ;
- ⤴ des outils disponibles dans des **bibliothèques open source**, Apache POI, CLibPST, Java LibPST, Java Mail, Mime4J, Tika, qui permettent une extraction des métadonnées depuis des messages et des messageries, ainsi qu'un traitement destiné à extraire en-tête, corps et pièces jointes et à identifier les formats de fichiers.

Les outils retenus par l'équipe-projet font l'objet de fiches descriptives en annexe 4. Les outils Aid4Mail, CERP Parser, EmailChemmy, Xena ont été téléchargés par un archiviste du Service historique de la Défense disposant de droits d'administration sur son poste de travail connecté au réseau internet. Les outils disponibles dans les bibliothèques *open source* ont quant à eux été installés et utilisés directement par le directeur technique du projet VITAM à la sous-direction des systèmes d'information du ministère de la Culture et de la Communication (MCC).

3.2. Résultats

3.2.1. Les outils

Les trois catégories d'outils utilisés pour les tests présentent des caractéristiques très différentes. Il ressort cependant de l'étude que les **outils les plus ergonomiques et les plus faciles d'utilisation ne sont pas ceux qui disposent des fonctionnalités les plus avancées.**

Aid4Mail et EmailChemmy

Outils du marché, Aid4Mail et EmailChemmy ont été installés et testés par le Service historique de la Défense (SHD) dans leur version de démonstration gratuitement disponible. Ils présentent de grandes similitudes, tant au niveau des fonctionnalités que de l'ergonomie générale.

Aid4Mail (voir annexe 4, outil 1) offre la possibilité à ses utilisateurs de traiter des messages enregistrés dans de très nombreux formats : formats d'exports des serveurs et clients de messagerie Microsoft (msg, pst) ; formats d'exports des principaux clients de messagerie (Apple Mail, Thunderbird, Windows Live Mail) ; formats standard (eml, mbox) ; formats exportés à partir de clients plus anciens (MSN Mail, Outlook Express v4) ou moins utilisés (Calypso, Eudora, Opera, Pegasus, Qualcomm). Il offre plusieurs fonctionnalités de traitement :

- ⤴ la conversion des messageries ou des messages vers des formats cibles : csv, eml, mbox, .PDF, pst, XML ;
- ⤴ l'extraction des pièces jointes et des carnets d'adresses ;
- ⤴ l'extraction des métadonnées aux formats suivants : csv, .tsv, XML ;
- ⤴ la réalisation d'un journal (*timeline*) par jour ou par mois¹⁷⁷.

Facile à prendre en main et à utiliser, d'une ergonomie agréable, cet outil présente néanmoins deux inconvénients majeurs, du moins dans la version de démonstration à la disposition de l'équipe-projet :

- ⤴ il ne permet de traiter les messages que par lot de 50¹⁷⁸ ;
- ⤴ pour le traitement des messageries exportées au format pst, il nécessite l'installation sur le poste de travail du client Outlook, ce qui n'était pas le cas sur le poste de travail utilisé au Service historique de la Défense pour la réalisation des tests.

EmailChemy (voir annexe 4, outil 4) offre également la possibilité de traiter des messages enregistrés dans de très nombreux formats : formats d'exports des serveurs et clients de messagerie Microsoft (pst) ; formats d'exports des principaux clients de messagerie (Apple Mail, Thunderbird, Windows Live Mail) ; formats standard (eml, mbox) ; formats exportés à partir de clients plus anciens (AOL) moins utilisés (Eudora). Il offre plusieurs fonctionnalités de traitement :

- ⤴ la conversion simple ou avancée des messageries ou des messages vers des formats-cibles, basés sur la RFC-2822 : csv, eml, mbox ;
- ⤴ l'extraction d'une seule catégorie de métadonnées, les adresses mails (*harvester*) ;
- ⤴ le découpage des fichiers mbox, soit par date, soit par taille (*splitter*)¹⁷⁹.

Aussi facile à prendre en main et à utiliser, même sans documentation technique, qu'Aid4Mail, il ne présente pas les mêmes inconvénients que celui-ci. Le Service historique de la Défense a pu traiter sans aucune difficulté non seulement le fichier pst utilisé pour les tests mais aussi sans limite de nombre de messages. En revanche, ses fonctionnalités sont moins riches que celles d'Aid4Mail, notamment en matière d'extraction des pièces jointes ou des métadonnées.

CERP Email Parser et Xena

Logiciels développés par des services d'archives dans le cadre des projets de pérennisation étudiés ci-dessus, CERP Email Parser et Xena ont tous deux pour objectif de traiter des messageries au format mbox en utilisant des enveloppes XML.

177 Voir le manuel utilisateurs disponible en ligne :

http://www.aid4mail.com/doc/Aid4Mail2_Manual.PDF (consulté le 22 juin 2013).

178 Il s'agit de la limite constatée sur l'outil testable en ligne.

179 Voir le manuel utilisateurs disponible en ligne :

http://www.weirdkid.com/products/emailchemy/doc/Emailchemy_User_Manual.PDF (consulté le 22 juin 2013).

Le CERP Email Parser (voir annexe 4, outil 3) maintenu par le Rockefeller Center et la Smithsonian Institution Archives a pour objectif d'aider les services d'archives à convertir des comptes de messagerie en fichiers XML qui contiennent les messages eux-mêmes avec leurs pièces jointes, en utilisant l'E-Mail Account Schema présenté ci-dessus. Il n'accepte en entrée que des messages enregistrés au format mbox.

Le logiciel peut être utilisé soit en utilisant une interface web, soit en utilisant l'interface native du logiciel Squeak¹⁸⁰. Son utilisation est cependant complexe et peu accessible pour toutes les catégories d'utilisateurs.

La conversion des fichiers mbox utilisés pour les tests au Service historique de la Défense (SHD) a échoué. À défaut de message d'erreur, il n'a pas été possible de diagnostiquer l'origine du problème. L'équipe-projet envisage de prendre contact avec les institutions concernées pour reprendre les tests ultérieurement.

Xena (voir annexe 4, outil 7), brique de la *Digital Preservation Software Platform* (DPSP) des Archives nationales d'Australie, est un logiciel libre qui offre deux fonctionnalités :

- ♣ l'identification du format des fichiers d'objets numériques ;
- ♣ la normalisation, la conversion et le stockage de ces objets dans un format de préservation propre. Pour les messageries électroniques, ils sont convertis en utilisant un paramétrage relativement souple (gestion des pièces jointes par exemple) dans un fichier XML avec création d'un index joint des fichiers au format mbox, pst et trim¹⁸¹.

Malheureusement, la conversion des fichiers mbox et pst utilisés pour les tests au Service historique de la Défense (SHD) a systématiquement échoué. Les messages d'erreurs générés par l'outil n'ont pas permis de définir si les dysfonctionnements du logiciel constatés étaient dus à un problème d'installation ou à un problème de normalisation des fichiers utilisés comme jeux de tests. L'équipe-projet a en conséquence pris contact avec les Archives nationales d'Australie pour diagnostiquer l'origine du problème.

Les outils des bibliothèques open source

Deux grandes catégories d'outils tirés des bibliothèques *open source* ont été utilisées : les outils de traitement des messages (Apache POI, CLibPST, Java LibPST, Java Mail, Mime4J, Tika) et les outils d'identification des formats de fichiers (DROID).

Pour le traitement des messages et l'extraction des informations ont été utilisées les bibliothèques suivantes :

- ♣ **Apache POI** (voir annexe 4, outil 2) : cette bibliothèque Java permet l'extraction de données des fichiers dont le format est propriétaire et en particulier des souches Microsoft (Word, Excel, PowerPoint, Outlook, ...). Compte-tenu de la faiblesse des champs nativement remontés dans l'interface de programmation de la bibliothèque, il a été nécessaire de manipuler les champs de manière programmée. Apache POI permet l'accès aux champs du message en mode chaîne de caractères, à partir desquelles il est possible d'extraire les

180 Voir le manuel utilisateurs disponible en ligne:

http://siarchives.si.edu/cerp/parser/EMAParserUserGuide_v1.PDF (consulté le 22 juin 2013).

181 Voir le manuel utilisateurs disponible en ligne: <http://xena.sourceforge.net/media/xenahelp.PDF> (consulté le 22 juin 2013).

informations souhaitées. Son usage n'est pas aisé et il n'y a pas de garantie sur les versions supportées. De plus, ce module n'est pas actuellement intégré dans la distribution officielle de POI. Il est donc nécessaire de procéder au téléchargement séparé de ce module pour pouvoir l'utiliser ;

- ⤴ **Java LibPST** (voir annexe 4, outil 5) et **CLibPST** : ces bibliothèques permettent de gérer des messageries au format pst (Outlook). Une grande partie des champs sont nativement accessibles grâce à l'interface de programmation de la bibliothèque. Cependant, pour être complet, mais aussi pour gérer les données retournées, des traitements programmatiques sont nécessaires. Ces bibliothèques permettent de récupérer nativement les corps des messages, aux formats HTML, rtf et/ou txt, en fonction de ceux qui sont disponibles dans le message. La différence entre les deux outils est le langage de programmation ;
- ⤴ **Java Mail** (voir annexe 4, outil 6) : bibliothèque Java éditée par Oracle, permet de gérer des messages (lecture, écriture, et même envoi) au format eml (norme RFC 5322). Son usage est assez simple au premier abord. La plupart des champs standards de la norme disposent de méthodes dans l'interface de programmation de la bibliothèque facilitant leur manipulation. Cependant, pour certains champs "peu standards" (comme le X-RDF) ou dans le cas de mauvaise lecture, des traitements programmés spécifiques doivent être intégrés. L'extraction des corps des messages nécessite une attention particulière, puisqu'il faut tenir compte du format sous-jacent, tant pour l'encodage du transport que pour l'encodage des caractères utilisé par l'outil de messagerie de l'émetteur. Cette bibliothèque ne traite que les messages au format eml. Pour traiter des messageries au format mbox, il est nécessaire de programmer l'outil afin qu'il analyse le fichier et le découpe en messages eml, avant de les passer à la bibliothèque Java Mail. Cette opération est complexe, car le format mbox n'est pas très stable de prime abord. Pour réaliser les tests, le ministère de la Culture et de la Communication s'est appuyé sur les travaux menés par la Queen's University, de Kingston (Canada) ;
- ⤴ **Mime4J d'Apache** : cette bibliothèque est censée gérer les messages aux formats eml et mbox, mais n'est en réalité pas très efficace et présente de sérieux problèmes dans le traitement des fichiers mbox ;
- ⤴ **Tika d'Apache** : cette bibliothèque permet la gestion des messages au format mbox ne peut extraire que le premier message de chaque fichier, ce qui rend son usage impossible. Elle a néanmoins été utilisée pour extraire les métadonnées des pièces jointes associées aux messages, notamment les mots-clés, car elle se révèle efficace pour l'extraction des textes de fichiers de multiples natures (XML, HTML, Microsoft Office, OpenDocument, PDF, rtf...).

Pour l'identification des formats de fichiers et l'extraction des métadonnées techniques ont été utilisées les bibliothèques suivantes :

- ⤴ **Java Fits** : cette bibliothèque permet d'extraire les métadonnées techniques des fichiers. Pour des raisons d'homogénéité, et surtout de contrôle de l'identification primaire des formats avec DROID, une surcouche a été réalisée pour appeler explicitement les bibliothèques JHOVE (1.9) et ExifTool (9.24). Enfin, une interface native a été réalisée pour l'utilisation de DROID (6.1). L'avantage de Fits est de proposer un premier niveau de schéma XML consolidé entre les différents outils.
- ⤴ **DROID 6.1** : bibliothèque développée par les Archives nationales du Royaume-Uni et s'appuyant sur les signatures du registre PRONOM V67 (26 février 2013), elle permet l'identification précise du format d'un fichier. L'outil a dû être reprogrammé sur quelques aspects, notamment pour prendre en compte les éléments suivants :

- ⤴ la gestion des archives TAR.BZIP2 (seuls les ZIP et TAR.GZIP sont pris en compte dans DROID 6.1) ;
- ⤴ la gestion des multiples formats d'identification, y compris les anciens formats. Ce développement permet de s'assurer qu'un fichier identifié avec une ancienne version de PRONOM soit toujours identifié correctement avec une nouvelle version. Par exemple, un document identifié comme un Word 2003 avec une première version de PRONOM, pourrait être identifié plus tard comme un Word 2007 avec une autre version de PRONOM. Avec la modification, l'ancien format est lui aussi retourné, ce qui permet de n'avoir qu'une alerte au lieu d'une erreur.
- ⤴ gestion des cas particuliers où DROID ne peut identifier un format de fichier, en particulier les formats eml et mbox. La modification permet d'appeler un parser eml ou mbox et ainsi de détecter les formats correspondants. L'identification du fichier est alors corrigée ;
- ⤴ **Jhove 1.9** : cette bibliothèque permet l'extraction de métadonnées techniques de fichiers ;
- ⤴ **Perl ExifTool** : cette bibliothèque permet comme JHOVE d'extraire des métadonnées techniques. Les deux logiciels ExifTool et JHOVE se sont avérés complémentaires en termes d'informations remontées, même si globalement ExifTool est plus pertinent.

Tous ces outils nécessitent donc des traitements programmatiques importants pour permettre leur utilisation optimale et la réalisation de traitements avancés. Leur usage est donc réservé à des professionnels.

3.2.2. Les traitements

Plusieurs traitements ont été réalisés sur les messageries de tests collectées par le Service historique de la Défense (SHD) et par le ministère de la Culture et de la Communication (MCC), en utilisant les fonctionnalités permises par les différents outils : conversion de format ; extraction des métadonnées ; extraction du corps et des pièces jointes ainsi que de mots-clés. Ces tests permettent de tirer un certain nombre de conclusions sur l'intérêt des différents outils ainsi que sur les traitements techniques possibles pour assurer une conservation à long terme des messageries électroniques.

Conversion d'un format vers un autre

La première partie des tests a concerné la conversion des fichiers collectés par le Service historique de la Défense aux formats eml, mbox et pst en utilisant Aid4Mail et EmailChemy, dans les limites expliquées ci-dessus pour le logiciel Aid4Mail. Les opérations de conversion des fichiers collectés n'ont pris à chaque fois que quelques secondes à quelques minutes.

Les tests suivants ont été réalisés et ont donné les résultats qualitatifs suivants :

Format d'origine	Format-cible	Logiciel	Résultat
eml	mbox	Aid4Mail	Création d'un fichier au format mbox dans la limite de lots

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

		(version de démonstration)	de 50 messages
		EmailChemy	Création d'un fichier au format mbox contenant l'ensemble des 545 messages d'origine. Échec de la conversion d'un message.
	.PDF	Aid4Mail (version de démonstration)	Création d'un fichier au format .PDF embarquant les pièces jointes
mbox	csv	EmailChemy	Création d'un fichier au format csv contenant l'ensemble des 369 messages, avec leurs pièces jointes
	eml	Aid4Mail (version de démonstration)	Conversion des 50 premiers messages
		EmailChemy	Création d'un dossier contenant l'ensemble des messages (369 dans le cas concerné) contenus dans le fichier d'origine (un seul dossier)
	.PDF	Aid4Mail (version de démonstration)	Création, à la demande de l'utilisateur, soit d'un seul fichier au format PDF contenant l'ensemble des messages et leurs pièces jointes, soit d'autant de fichiers .PDF que de messages.
pst	eml	EmailChemy	Échec de la conversion
	mbox	EmailChemy	Création d'autant de fichiers mbox que de dossiers et sous-dossiers présents dans le fichier d'origine

D'un point de vue général, les trois fichiers tests ont pu être convertis dans des formats basés sur la RFC 2822.

Les points suivants méritent cependant d'être signalés :

- ♣ les **possibilités de conversion du fichier pst ont été très limités** en raison du défaut d'installation des modules ou logiciels complémentaires nécessaires au bon fonctionnement des logiciels ;
- ♣ le **contrôle des formats eml et mbox n'a pas été possible via le logiciel DROID** des Archives nationales du Royaume-Uni (version 6.1), à défaut d'enregistrement dans le registre PRONOM d'une signature correspondant à ces formats. En revanche, ces formats ont été validés par le ministère de la Culture et de la Communication en utilisant le logiciel Java Mail ;
- ♣ les **fichiers au format .PDF sont au format PDF 1.4**, conformément à l'analyse effectuée avec le logiciel DROID des Archives nationales du Royaume-Uni (version 6.1). Format d'impression, le format .PDF 1.4 ne récupère qu'une partie des métadonnées de l'en-tête du message ;
- ♣ la **conversion de fichiers eml au format mbox ne permet pas la conservation de la mise en page du corps du texte.**

Extractions

La deuxième partie des tests a concerné l'extraction de tout ou parties des messages, en utilisant principalement les logiciels extraits des bibliothèques *open source*. Leur utilisation a nécessité une programmation par le ministère de la Culture et de la Communication (MCC). Les opérations d'extraction avec les logiciels ainsi reprogrammés n'ont pris que quelques minutes.

Des extractions de pièces jointes et de métadonnées ont également été réalisées par le Service historique de la Défense (SHD) en utilisant les logiciels Aid4Mail et EmailChemmy. Elles ont été réalisées en quelques minutes.

En-têtes, corps, pièces jointes et fils de discussion

Seuls Aid4Mail et les logiciels tirés des bibliothèques *open source* ont permis de réaliser des extractions plus ou moins complètes.

Le logiciel Aid4Mail a permis au Service historique de la Défense (SHD) d'extraire les pièces jointes d'un des messages collectés au format eml. Les sept documents, six au format png et un au format vsd, avaient été joints au message dans un dossier compressé. Le logiciel Aid4Mail a restitué le dossier compressé tel quel, accompagné d'un index au format .tab reprenant les informations suivantes : fichier ; date ; expéditeur ; sujet ; destinataires ; identifiant du message ; dossier ; dossier source (Folder Source) ; nom du contenu (ContentName). L'expérience a été renouvelée avec succès sur un lot de 50 messages, limite permise par le logiciel, du moins dans sa version de démonstration.

Les logiciels des bibliothèques *open source* permettent des extractions bien plus complètes, que ce soit pour les fichiers mbox (logiciel Java Mail) ou pour les fichiers pst (logiciel Java LibPST).

Pour le fichier mbox collecté par le Service historique de la Défense (SHD), le logiciel Java Mail a permis d'extraire les éléments suivants :

- ♣ un fichier main-mbox.xml constituant l'index des messages extraits et reprenant l'ensemble des informations présentes dans les en-têtes des messages. Conformément à la programmation réalisée, le fichier d'index a été enrichi avec l'analyse du format des fichiers extraits en utilisant le logiciel DROID des Archives nationales du Royaume-Uni (dans sa version 6-1), ainsi qu'avec les mots-clés extraits automatiquement (cf. ci-dessous) ;
- ♣ un dossier par message extrait. Chaque dossier contient le corps du message, aux formats HTML et/ou txt, ainsi que les pièces jointes au message (y compris les cartes de visites et les cartouches). Si un message avait pour pièce jointe un autre message, celui-ci fait l'objet d'un sous-dossier dans le dossier du message d'origine ;
- ♣ un fichier erreur.txt, indiquant les pièces jointes corrompues au moment de l'extraction. Il s'agit de pièces jointes dont le nommage comprend un caractère accentué. Les erreurs identifiées sont sans doute dues à un problème lors de la transmission des pièces. Celles-ci sont absentes des dossiers-cibles.

Un problème majeur a néanmoins été constaté à l'occasion de ces extractions. Le logiciel Java Mail ne décode pas convenablement les caractères accentués des en-têtes ou des corps des messages, que ceux-ci soient convertis au format .html ou au format txt.

Pour le fichier pst collecté par le SHD, le logiciel Java LibPST a permis d'extraire les éléments suivants :

- ♣ un fichier d'index général main.xml ;
- ♣ une arborescence de dossiers correspondant à l'arborescence de la boîte d'origine. Chaque dossier comprend les éléments suivants :
 - ♣ un fichier index info_nomdudossier.xml, comprenant l'ensemble des informations extraites de l'en-tête des messages contenus dans le dossier, les métadonnées extraites des propriétés des pièces jointes ainsi que les mots-clés extraits automatiquement (cf. ci-dessous). Le nom des utilisateurs est celui tiré de l'annuaire sur lequel s'appuie la messagerie. Les caractères accentués sont correctement restitués ;
 - ♣ un dossier par élément contenu dans le dossier correspondant, que ce soit un message ou un sous-dossier. Chaque dossier correspondant à un message comprend le corps du message aux formats .html, rtf ou txt, ainsi que les pièces jointes aux messages (y compris les cartes de visite et les cartouches). Les caractères accentués sont correctement restitués dans les fichiers aux formats rtf ou txt, mais pas dans les fichiers au format .html.
- ♣ un fichier erreur.txt, indiquant les pièces jointes corrompues au moment de l'extraction. Les erreurs identifiées sont sans doute dues à un problème lors de la transmission des pièces. Celles-ci sont néanmoins présentes dans les dossiers-cible, sans toutefois pouvoir être ouvertes.

Sur la base des informations extraites des métadonnées (Message-ID et In-Reply-To), les fils de discussion peuvent être reconstitués. Il est donc possible de reconstituer le parcours des messages en utilisant le fil de discussion, y compris dans un contexte d'archivage électronique.

Métadonnées

Tous les logiciels testés dans le cadre de la présente preuve de concept ont permis d'extraire des métadonnées des messages électroniques.

EmailChemy est le logiciel avec lequel les résultats sont les plus ténus. Seules les adresses des correspondants du titulaire de la boîte ont pu être extraites dans un fichier au format txt. Il convient de noter que ce fichier comprend également le contact de l'éditeur du logiciel ("Weird Kid Software" <emailchemydemo@weirdkid.com>).

Aid4Mail offre davantage de possibilités. Une fonctionnalité permet d'extraire les métadonnées suivantes aux formats csv, tsv et XML : date d'expédition ; expéditeur ; destinataires ; objet ; niveau de priorité ; identifiant du message ; corps du texte (sans les balises de mise en page du format HTML) ; nom des pièces jointes. Les métadonnées de traçabilité du message sont néanmoins absentes dans le fichier cible.

Comme pour les différents composants du message, **les logiciels tirés des bibliothèques open source fournissent les résultats les meilleurs**. L'ensemble des métadonnées, y compris les métadonnées de traçabilité, ont pu être extraites dans les différents fichiers XML indiqués ci-dessus. Le défaut d'encodage des caractères accentués constitue néanmoins un défaut non négligeable dans le contexte francophone, mais il n'est guère surprenant pour des outils développés dans un contexte anglo-saxon.

Mots-clés

Le ministère de la Culture et de la Communication (MCC) a programmé les logiciels tirés des bibliothèques *open source* pour extraire des mots-clés qui sont venus enrichir les index des messageries. L'objectif était de montrer que cette approche était faisable.

Cette indexation a consisté à extraire les mots les plus répétés dans les objets, les corps et les pièces jointes des messages. L'algorithme général d'extraction des mots clefs est assez naïf dans la version de la preuve de concept :

- ⤴ les mots du titre ont un poids maximum ;
- ⤴ les mots du corps du texte ont un poids relatif au nombre d'occurrences ;
- ⤴ les mots des pièces jointes ont un poids relatif au nombre d'occurrences divisé par la profondeur de jointure. Ainsi les pièces jointes de premier niveau ont un poids deux fois moindre que le corps du courriel. Dans le cas où une pièce jointe était elle-même un courriel, celui-ci a été également analysé (comme un courriel) et ses pièces jointes avaient alors un poids deux fois moindre que le courriel en pièce jointe (soit 4, 8, ... fois moins en fonction de l'imbrication) ;
- ⤴ seuls les 100 premiers mots-clés représentatifs (de poids le plus fort) ont été conservés.

Cette expérimentation pourra être améliorée à l'avenir en utilisant les outils techniques qui sont en pleine croissance et développement aujourd'hui, autour de la sémantisation des données. Pour le moment limitée au domaine du web sémantique, la sémantisation devrait permettre d'automatiser une partie des processus d'identification et de classement, sans remplacer néanmoins totalement l'intervention humaine.

En conclusion, on peut dire que **les extractions obtenues avec les logiciels tirés des bibliothèques *open source* sont donc considérablement plus riches que celles obtenues avec Aid4Mail**. Ils permettent la dissociation encadrée des différents composants d'une même boîte, ce qui rend possible l'établissement d'une stratégie d'archivage adaptée à chacun d'entre eux. Ils constituent donc une brique intéressante pour le traitement des messageries électroniques dans la perspective d'un archivage durable.

3.2.3. L'intégrité des exports

Objectifs et descriptifs des tests

L'objectif de ces tests complémentaires consistait à vérifier la capacité des outils informatiques actuels, et notamment les logiciels tirés des bibliothèques *open source*, à extraire des messages les informations (métadonnées) des messages :

- ⤴ RFC 5322 : champs adresses (From, Sender, Reply-To, To, Cc, Bcc) ; Subject ; date d'émission et date de réception ; traces des serveurs de distribution ; identifiants des messages et fils de discussion (Message-ID, In-Reply-To, References) ; niveau d'importance, de priorité et de sensibilité ; présence, nom et taille de pièces jointes ;
- ⤴ données spécifiques aux formats eml et mbox : dossier *via* le chemin et le nom du fichier

MBOX ; date de création de la pièce jointe ; champs RFC se rapportant au format MIME : Content-Encoding, Description, Content-Type, Content-Transfer-Encoding, Content-Language, ContentId, Content-Disposition ;

- ⤴ données spécifiques aux formats msg et pst : Folder, ConversationTopic, forwarded, replied, read, unsent, recipientReassignmentProhibited.

Les extractions de métadonnées ont ainsi pu être comparées entre les différents formats et les différents outils. L'extraction des informations s'est appuyée sur les normes suivantes :

- ⤴ RFC 5322 d'octobre 2008 (mise à jour de RFC 4021 mars 2005) commune aux formats pst/mbox et msg/eml ;
- ⤴ MIME RFC 2045 à 2049 (novembre 1996), 2231 (novembre 1997), 3676 (février 2004), 4288, 4289 (décembre 2005), 5147 (avril 2008), 6532, 6533 (février 2012), 6657 (juillet 2012) pour les formats mbox et eml ;
- ⤴ Les spécifications du format pst (telles que décrites dans les outils de manipulation de ce format) ;
- ⤴ RFC 4155 de septembre 2005 pour le format mbox.

Résultats

Les tests menés ont été concluants, à savoir qu'ont pu être extraites les informations (métadonnées) des messages, ainsi que les autres propriétés.

Format	Métadonnées RFC 5322	Autres propriétés	Bilan et observations
eml	oui	oui	L'extraction des métadonnées est très complète et très efficace.
mbox	oui	oui	L'extraction des métadonnées est très complète et très efficace.
msg	partiellement	partiellement	<p>Certaines métadonnées ne peuvent être extraites sans programmation de l'outil :</p> <ul style="list-style-type: none"> ⤴ conversationTopic ; ⤴ traces de réception ; ⤴ taille du courriel (en octet) ; ⤴ importance ou priorité du courriel ; ⤴ sensibilité du message. <p>Certaines métadonnées ne peuvent être récupérées, en raison de leur absence dans le fichier (mais présentes dans un export au format pst) :</p> <ul style="list-style-type: none"> ⤴ propriétés liées à la gestion des courriels dans Outlook : forwarded, replied, read, unsent, recipientReassignmentProhibited ; ⤴ la date de création d'une pièce jointe. <p>Les adresses sont au format Outlook, avec utilisation d'une notation de type domaine : Susan </O=ENRON/OU=NA/CN=RECIPIENTS/CN=SBAIL</p>

			E2>. Cette notation dépend d'un élément extérieur, à savoir le contrôleur de domaine auquel est rattachée l'adresse du courriel. De nombreuses erreurs ont été constatées sur le formatage correct des adresses, voire sur leur reconnaissance.
nsf	-	-	À noter qu'il existe une procédure pour exporter depuis Lotus Notes les messages directement aux formats eml ou mbox.
pst	partiellement	oui	Les métadonnées importantes ont été extraites avec succès, dans une qualité suffisante. L'absence d'extraction de métadonnées est due à leur absence dans le fichier correspondant : <ul style="list-style-type: none"> ⤴ les champs Sender et Reply-To ne sont pas renseignés ; ⤴ les champs de formatage sont gérés de manière spécifique dans le format pst ; ⤴ le champ Reference est absent.

Les résultats détaillés des tests d'extraction font l'objet d'une synthèse en annexe 5.

Le cas des messageries officielles

La vérification de l'intégrité des messages extraits de messageries officielles constituait le complément nécessaire aux tests réalisés précédemment. Aussi plusieurs messages collectés depuis la Messagerie Melind@ ont-ils fait l'objet d'une analyse utilisant les résultats de l'étude InSPECT.

Il ressort de cette analyse que les messages au format eml extraits de Melind@ sont en tous points identiques à ceux extraits de messageries interpersonnelles. Les données relatives à la rédaction, à la validation et à la signature du message sont en fait intégrées dans le corps même du message, rédigé au format .html, sous la forme d'un tableau qui comprend les blocs suivants :

- ⤴ à gauche, un bloc indiquant le nom et les coordonnées du rédacteur ;
- ⤴ à droite, en haut, un bloc indiquant le nom de l'entité émettrice, la date de l'émission et le niveau de classification de l'ordre ;
- ⤴ à droite, au milieu, un bloc indiquant les entités destinataires de l'ordre ;
- ⤴ à droite, en bas, un bloc indiquant l'objet de l'ordre, la référence à un ordre précédent, ainsi que la présence et le nom des pièces jointes.

Toutes les informations d'intégrité identifiées par l'étude InSPECT dans le scénario idéal sont conservées dans les exports au format eml. Les opérations de conversion et d'extraction des métadonnées ont donc pu être réalisées sans aucune difficulté. La conservation à long terme de ces messages ne pose donc aucune difficulté particulière.

3.3. Conclusions

Les tests réalisés par l'équipe-projet VITAM permettent de compléter les conclusions tirées par les précédentes études.

3.3.1. Bilan par outil

La première conclusion est relative aux outils utilisés pour réaliser l'ensemble des tests.

Premièrement, l'étude confirme que des **solutions techniques existent pour la conversion et l'extraction des informations des messageries électroniques**. Certes, les outils disponibles ne donnent pas tous entièrement satisfaction, que ce soit au niveau de l'ergonomie ou des fonctionnalités disponibles et surtout de l'encodage des jeux de caractères. Cependant, les résultats obtenus sont encourageants.

Deuxièmement, les **outils développés dans le cadre des projets de coopération par les services d'archives publics se sont révélés les moins facilement utilisables**. Si leur installation n'a pas posé de problème particulier en raison de la documentation technique disponible, leur utilisation pour l'archivage des messageries électroniques ne semble à ce jour pas une piste sérieuse à envisager. Les contacts pris avec les équipes chargées de leur maintenance permettront peut-être de mieux réviser ce premier jugement.

Troisièmement, les tests réalisés à partir du fichier pst collecté par le SHD montrent **l'importance de la plate-forme d'installation des différents outils**. L'absence du client de messagerie Outlook ou du module complémentaire Readpst sur le poste de travail utilisé a empêché l'équipe-projet de réaliser tous les tests souhaités.

Enfin, les **outils tirés des bibliothèques open source constituent la piste la plus sérieuse à étudier en vue de traiter les messageries électroniques**. C'est vraisemblablement autour de leur utilisation que l'équipe-projet VITAM va poursuivre ses investigations afin de fournir une réponse technique à la conservation des messageries électroniques dans le contexte français.

Le tableau qui suit résume les principales conclusions de l'équipe-projet sur les différents outils utilisés :

Outil	Formats en entrée	Commentaires
Aid4Mail	eml, mbox, msg, pst	Ergonomique, simple d'utilisation. Nombreuses fonctionnalités : conversion, extraction des métadonnées et des pièces jointes. Impossible à utiliser pour la conversion des formats pst sans l'installation d'Outlook sur le poste de travail. Ne permet le traitement que de lots de 50 messages, du moins dans la version de démonstration.
Apache POI	msg	Permet d'accéder au contenu des en-têtes, et donc aux propriétés additionnelles, de manière manuelle. POI-HSMF destiné à la gestion des fichiers msg Nécessite des traitements programmatiques

Auxilii MsgParser	msg	Cette bibliothèque permet de gérer des messageries au format msg . Elle ne comporte aucune fonctionnalité supplémentaire de traitement des en-têtes des courriels par rapport à Apache POI (version standard) sur lequel elle s'appuie.
C LibPST	pst	Nécessite des traitements programmatiques
CERP Email Parser	mbox	Échec de l'utilisation Contacts à prendre avec le Rockefeller Center et la Smithsonian Institution Archive
EmailChemy	eml, mbox, msg, pst	Ergonomique, simple d'utilisation. Plusieurs fonctionnalités : conversion, export vers Gmail, moissonnage de métadonnées, découpage de fichiers mbox.
Java LibPST	pst	Nécessite des traitements programmatiques
Java Mail	eml, mbox	Nécessite des traitements programmatiques
Mime4J	mbox	Usage quasiment impossible => outil abandonné
Tika	mbox	Ne permet d'extraire que le premier message. Outil abandonné.
Xena	mbox	Ergonomique et simple d'utilisation. Échec des tests. Contacts pris par l'équipe-projet avec les Archives nationales d'Australie.

3.3.2. Bilan par format

La seconde conclusion est relative aux différents formats d'export des messageries électroniques.

Sans surprise, les tests réalisés dans le cadre de la présente preuve de concept confirment les résultats obtenus par les précédentes études et notamment l'intérêt des formats eml, mbox et pst, avec quelques remarques et réserves exposées ci-dessous pour les deux derniers formats.

Le format mbox, défini par la RFC 4155, stocke les messages au format natif eml dans un seul fichier, en utilisant comme séparateur une ligne de texte définie comme suit :

- ♣ au début : « From optional-email-address timestamp »
- ♣ en fin : une ligne vide.

Ce format impose le parcours intégral du fichier pour identifier les différents messages. C'est la raison pour laquelle de nombreux clients de messageries utilisent un fichier d'indexation externe au fichier mbox pour gérer les emplacements des messages de façon unitaire.

Le format pst, bien que propriétaire, est un candidat sérieux pour la collecte des messageries, non seulement parce qu'il constitue le format d'export du client de messagerie le plus répandu jusqu'à peu, mais aussi parce qu'il permet de récupérer toutes les données importantes avec une qualité suffisante.

Il faut cependant noter que le risque d'obsolescence de ce format est très important. Un fichier exporté à partir du logiciel Outlook 2003 ne pourra par exemple pas obligatoirement être relu dès lors que l'utilisateur aura migré vers un environnement technique utilisant le logiciel Outlook 2010. Aussi, si ce format est acceptable en entrée pour l'archivage, ne peut-il pas l'être pour la conservation. Il convient en conséquence de traiter le plus tôt possible les fichiers exportés à ce format afin d'éviter les difficultés de relecture liées à cette obsolescence.

Le tableau qui suit résume les principales conclusions de l'équipe-projet sur les différents formats d'exports analysés :

Format	Format valable pour la constitution des SIP	Format valable pour la constitution des AIP	ID PRONOM	Commentaires
eml	Oui	Non	fmt/278	<p>Ce format est un très bon candidat pour l'entrée de messages unitaires.</p> <p>Il pourrait être considéré comme un format pérenne, si trois réserves ne devaient être faites :</p> <ul style="list-style-type: none"> △ les pièces jointes doivent pouvoir être converties pour être intelligibles et accessibles dans le temps, en prenant en compte le problème de l'encodage des caractères ; △ les métadonnées relatives au classement initial du message dans un dossier local sont perdues ; △ la conservation des métadonnées relatives au lien entre messages (fils de discussion) dépend de la manière dont l'export a été réalisé.
mbox	Oui	Non	-	<p>Ce format est le candidat idéal pour une entrée de comptes de messageries car il est facile d'accès, y compris <i>via</i> une lecture humaine non automatisée (recherches manuelles possibles). Il présente l'avantage de conserver les liens entre courriels, ainsi que le plan de classement défini par l'utilisateur.</p> <p>Il existe en cinq variantes, selon les outils de messagerie utilisés.</p> <p>Il ne peut cependant pas être considéré comme un bon candidat pour la conservation, car :</p> <ul style="list-style-type: none"> △ les pièces jointes nécessitent d'être converties pour être intelligibles et accessibles dans le temps, en prenant en compte le problème de l'encodage des

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

				<p>caractères ;</p> <p>^ l'accès unitaire à un courriel parmi d'autres est compliqué, étant donné qu'il faut parcourir l'ensemble du fichier pour retrouver ou extraire un message précis. En effet, les balises de fin de message peuvent générer des erreurs d'interprétation et on ne dispose pas toujours de fichier d'indexation.</p>
msg	Non	Non	x-fmt/430	Format binaire non ouvert, mal documenté et donc mal géré par les outils Open Source. Ce format est fortement déconseillé.
.nsf	-	-	x-fmt/336 x-fmt/337 x-fmt/338	Format non ouvert.
pst	Oui	Non	x-fmt/248 (1997-2002) x-fmt/249 (2003-2007)	<p>Format non ouvert.</p> <p>Ce format est un bon candidat pour une entrée de comptes de messageries</p> <p>Il existe un risque qu'un fichier enregistré dans une version ancienne du format ne soit plus supporté par des outils plus récents.</p> <p>Les extractions et conversions doivent être faites le plus tôt possible après la collecte, afin d'éviter de devoir faire face à de grandes difficultés de relecture pour des raisons d'obsolescence.</p>

Conclusion et propositions

Au terme de cette étude, l'équipe-projet VITAM ne peut que tirer une seule et unique conclusion : la conservation à long terme des messageries électroniques est possible et de nombreuses pistes techniques ont été explorées au niveau international.

Il apparaît clairement que les **véritables enjeux de cette conservation ne sont pas techniques, mais organisationnels, juridiques et archivistiques, notamment pour les messageries interpersonnelles**. Le caractère personnel de l'utilisation des messageries électroniques interpersonnelles et l'importance des usages font de chaque messagerie un cas particulier à traiter. La place centrale occupée par les messageries au sein du système d'information des organisations ne rend le sujet que plus crucial.

Propositions générales

Les rédacteurs de la présente étude sont amenés à émettre plusieurs propositions d'ordre organisationnel, juridique, technique et archivistique. Elles ont été soumises au Comité interministériel aux archives de France (CIAF) le 1^{er} juillet 2013 pour faire l'objet d'une discussion entre l'administration des archives, l'administration des systèmes d'informations et les représentants des producteurs d'archives publiques.

Propositions organisationnelles

Les usagers sont au cœur des problématiques liées à la conservation des messageries électroniques. C'est donc à eux que les principales actions doivent être destinées.

Afin de parvenir à une conservation efficace des décisions et traces de l'action administrative, **quelques règles simples relatives à l'usage des messageries et à la nature des informations échangées avec cet outil doivent être rappelés au plus haut niveau**, notamment aux hommes politiques, aux hauts fonctionnaires ou aux membres des cabinets ministériels qui, au quotidien, placent les messageries électroniques au cœur de leur processus de prise de décision :

- ♣ caractère engageant juridiquement des décisions et avis pris dans des messages ;
- ♣ nature d'archives publiques des messages produits ou reçus dans le cadre de l'activité professionnelle et obligation de les exporter pour remise à l'administration des archives lors de la cessation de fonction ;
- ♣ nécessité de séparer les messages relevant des activités professionnelles des autres messages (vie privée, activités associatives et politiques) ;
- ♣ risques à utiliser des outils non recommandés (appareils personnels, *webmails* du marché type Gmail, Apple mail ou Yahoo mail) pour s'assurer de la sécurité des systèmes d'information.

À l'égard de cette catégorie d'utilisateurs, seules des actions de sensibilisation courtes mais percutantes peuvent être envisagées, à l'image de celles proposées par le Club des responsables de politiques et projets d'archivage (CR2PA), notamment sous la forme de vidéos.

Des actions de sensibilisation plus classiques peuvent être envisagées pour les autres utilisateurs de la sphère publique : campagne de communication ; rédaction d'un guide de bonnes pratiques de l'usage des messageries, s'inspirant des guides déjà existants ; mise en place de formations à l'usage des messageries moins orientées sur les aspects techniques que sur les aspects de bonne gouvernance ; sensibilisation des utilisateurs lors de leur prise de fonction ; définition et diffusion de procédures de collecte des messageries au départ des agents (mutation, fin de contrat, départ en retraite, décès). Les cellules informatiques chargées de l'installation des postes informatiques dans les différentes administrations, qui sont directement en contact avec les utilisateurs, constituent des relais naturels de cette sensibilisation. Des actions ciblées de sensibilisation et de dialogue avec ces entités constitueraient sans doute une piste de travail intéressante.

Le recours à des messageries fonctionnelles ou officielles ne peut également qu'être encouragé au sein de l'administration. De nombreux ministères ont déjà mis en place des messageries officielles héritières des anciens modes de transmission (ministère des Affaires étrangères ou ministère de la Défense) ou des comptes de messageries fonctionnelles. Cette pratique devrait se développer si était adopté le projet de loi autorisant le gouvernement à rendre opposables par ordonnance les messages électroniques adressés par les citoyens à l'administration¹⁸². En effet, en n'étant pas liées à un utilisateur individuel en tant que tel, non seulement elles donnent un caractère officiel aux informations échangées par le biais de l'outil mais aussi elles évitent la présence de messages couverts par le respect de la vie privée. L'archivage de ces messageries ne pose dès lors aucun problème juridique préalable.

Une collaboration entre les administrations des archives, des systèmes d'information et de la modernisation de l'État permettrait sans aucun doute d'initier les premières actions.

Propositions juridiques

Le chantier juridique est sans doute le plus délicat à mener mais aussi le plus urgent au regard des points soulevés dans la partie 1.2. de la présente étude. Il nécessite une collaboration entre l'administration des archives, les services juridiques ministériels et interministériels et la Commission nationale de l'informatique et des libertés (CNIL), pour traiter des problèmes liés à la cohabitation entre données publiques et données privées dans les usages de l'outil.

La **rédaction de clauses-types dans les règlements intérieurs et les chartes informatiques** relatives à la gestion du cycle de vie et à l'archivage des messageries utilisées par les agents publics constitue une première piste à explorer. Les clauses à définir doivent permettre de clarifier les procédures imposées par les employeurs publics à leurs collaborateurs pour l'archivage des messages qu'ils produisent et reçoivent dans le cadre de leurs activités, dans le respect des principes de transparence et de proportionnalité imposés par la réglementation en vigueur. Une discussion avec les organisations représentatives du personnel sera naturellement nécessaire. Des dispositions spécifiques doivent également être prises pour les personnes mentionnées à l'article L 213-4 du code du patrimoine. Il conviendra sans doute de vérifier que les circulaires diffusées par le Secrétariat

182 Voir le dossier législatif publié sur le site du Sénat à l'adresse suivante : <http://www.senat.fr/dossier-legislatif/pj112-664.html> (consulté le 23 juin 2013).

général du gouvernement (SGG) et les dispositions des protocoles signés en application de l'article du code du patrimoine précité suffisent à garantir les principes de transparence et de proportionnalité.

L'élaboration d'un protocole-type à faire signer par l'agent au moment de la collecte de sa messagerie par un service d'archives publiques pourrait constituer une piste d'action complémentaire. Elle n'est cependant pas recommandable et ne doit être utilisée que de manière exceptionnelle, en raison de sa lourdeur administrative. Les exemples utilisés par la Mission des archives placée auprès du ministère de la Santé lors de l'archivage des cabinets ministériels au printemps 2012 ainsi que par la Mission des archives du ministère de la Culture et de la Communication lors de la dissolution du Groupement d'intérêt public France Télé Numérique pourraient servir de bases de travail intéressantes et rapidement mobilisables pour rédiger un protocole type.

Propositions techniques

Même si les aspects techniques ne constituent pas l'aspect aujourd'hui le plus crucial de la conservation des messageries électroniques, plusieurs actions pourraient également être initiées dans ce domaine.

La diffusion aux services informatiques de consignes inspirées des recommandations du groupe de travail InterPARES pourrait constituer une première piste de réflexion. Pour s'assurer de la valeur juridique des messages, leurs études démontrent que l'identification et le rattachement hiérarchique des expéditeurs au sein des organismes sont indispensables pour assurer la valeur juridique des messages. La tenue d'un registre auditable liant chaque adresse d'expédition à une personne physique (nom, prénom, entité de rattachement) ou morale (nom de l'entité, entité de rattachement) clairement identifiable et permettant au destinataire de connaître sans aucune ambiguïté possible l'identité de son correspondant fait notamment partie de ces bonnes pratiques recommandées¹⁸³. Pour les administrations de l'Etat, le chantier actuellement menée par la Direction interministérielle des systèmes d'information et de communication de l'Etat (DISIC) sur la gouvernance des données de référence constitue sans doute une opportunité pour vérifier la conformité des annuaires ministériels à ces recommandations. Cette proposition ne résout cependant en la question de l'identification des correspondants externes à l'organisme, problématique qui va devenir essentielle lorsque les messages adressés par les citoyens à l'administration vont lui devenir pleinement opposables.

Le développement de modules complémentaires, greffés notamment au client Thunderbird, et permettant l'export des messages importants soit dans un serveur de fichiers, soit dans une GED ou un ERMS, sur le modèle du projet LiveLink conduit par la République et Canton de Genève, constitue une deuxième piste intéressante. Au regard des observations effectuées par l'équipe-projet VITAM, une attention particulière devra être apportée à l'encodage des caractères accentués.

183 The InterPARES 3 Project, TEAM Italy, *Op. cit.*, p. 37-38, recommandation 3.2.10. On peut citer à titre d'exemple le registre utilisé par le ministère de l'Écologie où l'expéditeur est systématiquement libellé comme suit : NOM Prénom - Structure de rattachement [prenom.nom@domaine.gouv.fr]. Exemple : NOM Prénom - Entité [nom.prenom@developpement-durable.gouv.fr].

Propositions archivistiques

Mais c'est sans doute la réflexion professionnelle au sein de la communauté archivistique qui doit faire l'objet des principales actions à mettre en œuvre dans les prochains mois. Les modalités de collecte, de traitement et de conservation des messageries doivent faire l'objet d'analyses plus approfondies que celles initiées dans le cadre de la présente étude.

La connaissance des pratiques des utilisateurs, prenant en compte les apports en matière de sociologie du travail et d'utilisation des technologies de l'information et de la communication dans la vie professionnelle, doit être largement approfondie pour **définir une stratégie de collecte des messageries électroniques adaptée aux usages des différentes catégories d'administrations et de producteurs**. Le tableau élaboré par les rédacteurs de la présente étude en 1.3., transformé en questionnaire dans le cadre d'une visite menée conjointement avec la Mission des archives de France auprès des services du Premier ministre au Secrétariat général pour les affaires européennes (SGAE), pourrait servir de base à la réalisation d'une enquête plus poussée. Une collaboration avec les administrations de la fonction publique et du travail ainsi que des laboratoires de recherche universitaires pourrait sans aucun doute être initiée. Il est d'ailleurs intéressant de constater que l'importance des usages est à ce jour également au cœur des réflexions des archivistes de Nouvelles-Galles du Sud¹⁸⁴.

En matière de collecte des archives, **l'identification de comptes de messageries-cibles à collecter dans chaque organisation**, piste suggérée par la direction des Archives de France en 2009, semble aux rédacteurs de la présente étude la piste la plus sérieuse à envisager. Cette collecte pourrait être faite sur plusieurs bases : niveau hiérarchique ; insertion dans un réseau de correspondants cible (groupe de travail, réseau d'experts) ; participation à l'exercice de fonctions vitales ou critiques pour l'organisation ; etc. Les Archives nationales des États-Unis, dans le cadre du projet Capstone, semblent d'ailleurs mener en ce moment même une approche similaire¹⁸⁵. Cette approche plus globale de l'archivage des messageries électroniques semble également pertinente pour la collecte des archives privées. La rédaction de conseils pratiques à destination des archivistes chargés de cette collecte pourrait faire l'objet d'une collaboration avec les autres institutions patrimoniales, notamment les bibliothèques, intéressées par cette question.

Le tri, le classement et la description des messageries collectées isolément ou dans le cadre de collectes plus globales de fonds d'archives publiques comme privées constituent un sujet majeur sur lequel les réflexions et les expérimentations sont encore relativement peu nombreuses. Le reclassement d'une messagerie électronique collectée est-il possible, comme les archivistes le pratiquent pour les archives papier ? Comment trier, classer et décrire un fond d'archives mixte comprenant à la fois des documents papier et des documents électroniques, des messageries électroniques et des documents bureautiques, des messageries interpersonnelles comme des messageries fonctionnelles ? Le coût de ces opérations est-il soutenable dans le contexte actuel ? La pratique archivistique ne peut qu'évoluer en la matière, eu égard à la masse croissante d'information

184 Voir par exemple le *post* publié le 30 mai 2013 par les Archives de Nouvelles-Galles du Sud et disponible à l'adresse suivante : <http://futureproof.records.nsw.gov.au/email-messages-have-corporate-value-to-organisations-but-are-still-tricky-to-manage-latest-state-records-survey-results/> (consulté le 23 juin 2013).

185 Projet de NARA Bulletin en date du 30 mai 2013, disponible à en ligne : <http://blogs.archives.gov/records-express/files/2013/06/Capstone-Email-DRAFT-NARA1.PDF> (consulté le 23 juin 2013).

à traiter à la suite de pareilles collectes. La communication présentée par la Mission des archives de France auprès du ministère de la Santé à l'occasion du Forum 2013 de l'Association des archivistes français (AAF) ne laisse aucun doute sur le sujet. **L'administration des archives pourrait proposer aux masters professionnels d'archivistique des différentes universités de lancer des sujets de recherche ou de stage sur ces problématiques de tri, de collecte et de traitement des fonds mixtes.** Enfin, **l'expérimentation d'outils d'analyse sémantique pourrait également constituer un axe de réflexion intéressant.** Des recherches fondamentales pourraient également être menées en partenariat avec des laboratoires de recherches sur les outils sémantiques.

Les modalités de conservation et de restitution à long terme des messageries électroniques collectées et traitées doivent également faire l'objet de travaux spécifiques. **L'élaboration d'un schéma XML de représentation des messageries**, tirant parti des travaux déjà menés aux Pays-Bas, en Australie et aux États-Unis, pourrait constituer une première action que l'administration des archives pourrait initier. Les différents ministères partenaires du projet VITAM, devant prendre rapidement en charge des messageries électroniques, pourraient mener ce chantier conjointement avec l'administration des archives.

Enfin, la **diffusion en français des travaux menés dans le cadre des groupes InterPARES et InSPECT** serait sans doute utile.

Pour mener ces différents chantiers, des **collaborations pourraient être initiées avec les autres services d'archives menant au niveau international** des recherches parallèles sur le sujet. Les contacts pris par les rédacteurs de la présente étude laissent entendre que les Archives de l'État en Belgique pourraient s'associer à la conception d'un schéma international XML de représentation des messageries et que les collègues anglais et américains seraient également partisans d'une réflexion internationale sur le sujet.

Propositions dans le cadre du programme VITAM

À l'issue de la revue de littérature et des tests réalisés, et en tenant compte de l'état de ses réflexions en matière de modélisation des métadonnées, l'équipe-projet VITAM a commencé à définir une méthodologie de collecte des comptes de messageries et des messages électroniques.

Son objectif est de :

- Garantir l'authenticité des contenus, en s'assurant qu'une version non modifiée des objets collectés, est bien conservée dans le système ;
- Faciliter la mise en œuvre d'opérations de conservation adaptées à chaque catégorie de fichiers, notamment pour les pièces jointes ;
- Permettre la restitution de tout ou partie des messages conservés, en fonction des requêtes des utilisateurs ;
- Rendre possible à terme le dédoublement des fichiers numériques stockés dans le système, afin d'optimiser les conditions de stockage.

Collecte des messageries

La définition de modalités de collecte des messageries électroniques constitue une priorité

pour les trois ministères partenaires du programme, que ce soit en matière de formats d'exports ou de modalités de traitement.

Les tests réalisés conduisent l'équipe-projet à recommander des extractions des messageries aux formats eml, mbox et pst. Si le format eml est parfaitement adapté aux messageries officielles, à la structuration très simple par émetteur et par date, il est sans doute moins adapté pour les messageries interpersonnelles gérées par les utilisateurs au moyen d'arborescences plus ou moins complexes. Dans ces cas, les formats mbox et pst sont manifestement préférables, car ils permettent d'exporter non seulement les messages mais aussi l'arborescence des dossiers. **S'agissant des comptes exportés au format pst, le délai de traitement entre l'export et le versement au service d'archives public compétent devra être le plus court possible, afin d'éviter les problèmes liés à la rapide obsolescence du format.** Des tests complémentaires seront nécessaires pour mieux spécifier les modalités de traitement des fichiers collectés à ce format.

La fourniture de schémas d'extraction et de moyens de traitements sécurisés des messageries (conversion et extraction de données et de métadonnées) aux Missions des Archives de France ayant déjà collecté des messageries au format pst constitue notamment une priorité absolue pour les Archives nationales. Étant donné la complexité de l'utilisation des bibliothèques *open source*, des développements informatiques seront manifestement nécessaires.

Constitution des SIP

Dans une plate-forme s'appuyant sur la solution VITAM, les *Submission Information Packages* (SIP) correspondant à des comptes de messagerie pourraient prendre la forme d'un fichier zippé comprenant :

- ♣ une arborescence de répertoires avec, pour chaque message, un répertoire contenant :
 - ♣ le corps du message, converti dans un format permettant d'assurer sa lisibilité et son intelligibilité (mise en page) dans le temps : format .txt pour les corps les plus simples ; format .html pour les corps disposant de mises en page plus abouties ;
 - ♣ les pièces jointes, à la fois dans leur format binaire d'origine et dans un format garantissant leur pérennisation, en fonction des normes et standards en vigueur. S'agissant des messages insérés comme pièces jointes, ils sont traités comme des messages en tant que tels (sous-dossier dans le dossier) ;
 - ♣ le ou les messages dans leur format original de collecte ;
- ♣ un fichier au format XML indiquant la liste des répertoires et fichiers faisant l'objet de la soumission (identifiant, taille, empreinte, formats) et donnant pour chaque répertoire, les métadonnées archivistiques - règles d'accès, sort final, durée de conservation -, descriptives - sans doute avec une structuration permettant d'isoler les partenaires (expéditeur, destinataires), le processus (priorité, importance, confidentialité), le thème (objet, mots-clés) et le contenu (corps, pièces jointes) de l'échange - et techniques nécessaires à leur gestion - notamment les métadonnées extraites de l'en-tête du message.

Le fichier .zip serait transféré conformément aux prescriptions de la future norme NF Z 44022 relative à la modélisation des échanges de données pour l'archivage (MEDonA).

Ces hypothèses de travail supposent l'existence, en amont du transfert dans la plate-forme,

d'un outil permettant de traiter les messageries collectées et de formater les SIP dans la structure retenue.

Le schéma ci-dessous résume cette hypothèse de travail.

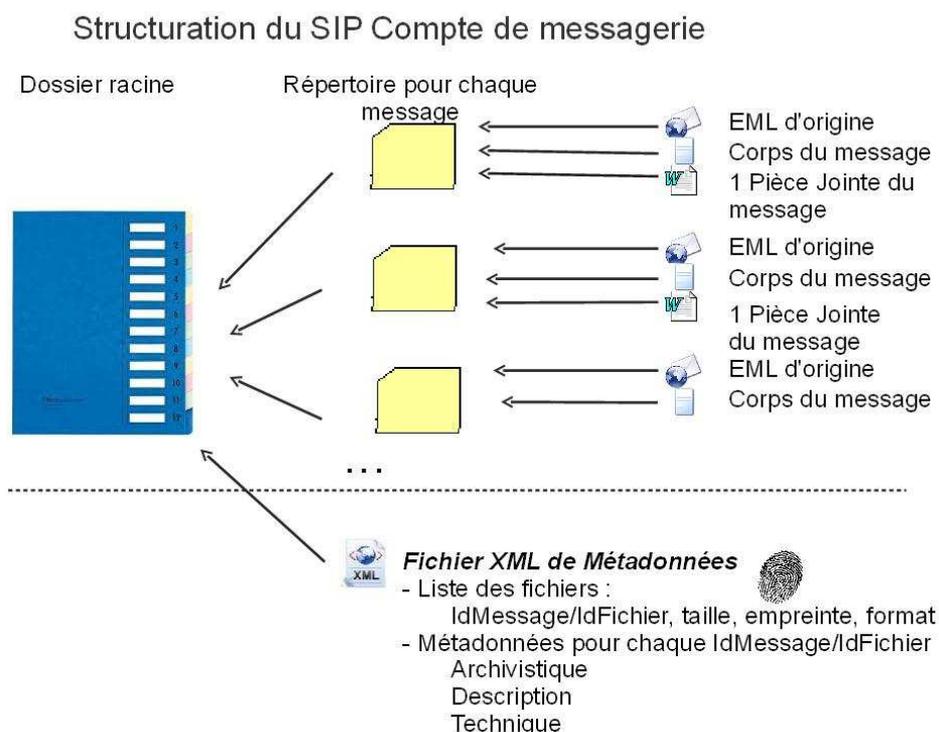


Fig. 7 : projet de schéma de représentation d'un SIP Compte de messagerie dans VITAM

Cette hypothèse sera consolidée par l'équipe-projet VITAM dans les prochains mois en poursuivant la réflexion autour du traitement des pièces jointes correspondant à des messages électroniques, en échangeant avec les services d'archives d'autres pays sur la problématique de l'encodage des caractères et en confrontant ces premières hypothèses aux exigences d'archivage de la messagerie officielle interarmée NeMO que le Service historique de la Défense doit définir d'ici la fin de l'année 2013.

Annexe 1 : Bibliographie

ADGENT Nancy. *Collaborative Electronic Records Project : An Introduction and Overview*. Sleepy Hollow, New York et Washington : The CERP Project, 2008, 29 p.

AGENCE NATIONALE POUR L'AMELIORATION DES CONDITIONS DE TRAVAIL (ANACT). *NTIC, de quoi parle t-on ?, Étude TICO : Technologies de l'Information et de la Communication dans les Organisations*, septembre 1999. 42 p., consultable en ligne: http://recherche.anact.fr/index.php?mid=6802109beac659551b90e915c94a1f51&l=fr&sortBy=pertinence&rub=on&searchOn=all&showExtr=1&offsetCooc=0&r34=1sr+cne4%5D&onlink=1&sx=1365586712_0&q=messageries+%E9lectroniques&offsetDisplay=0&collection=10&typeQuery=4 (consulté le 9 avril 2013).

ALQUIER Éléonore, BENOIST Marie-Élodie, LHOUMEAU Hélène, « Évaluer, sélectionner et collecter les documents numériques non structurés : l'exemple de la collecte 2012 dans les cabinets ministériels (Affaires sociales) », Présentation effectuée au Forum des archivistes de l'Association des archivistes français (AAF), 22 mars 2013.

ANDERSEN Kim N., MEDAGLIA Rony, VATRAPU Ravi, HENRIKSEN Helle Z., GAULD Robin. « The forgotten promise of e-government maturity: assessing responsiveness in the digital public sector », dans *Government information quarterly*, octobre 2011, vol. 28, n°4, p. 439-445.

ARCHIVES NATIONALES. *La gestion et l'archivage des courriels, manuel pratique*, version 2, août 2012, 60 p.

ARNOLD M. « On the phenomenology of technology : the « Janus-faces » of mobile phones », dans *Information and Organization*, 2003, n°13, p. 231-25.

ASSADI H., DENIS, J. « Les usages de l'e-mail en entreprise : efficacité dans le travail ou surcharge informationnelle ? », dans KEESOUS E. et METZGER J.-L., *Le travail dans les technologies de l'information*. Paris : Lavoisier, collection technique et scientifique des télécommunications, 2005, p. 136-155.

ASSOCIATION DES ARCHIVISTES FRANCAIS (AAF). *L'archivage des emails : quelles étapes ?*, présentation au Groupe PIN du 17 janvier 2011, consultable en ligne: http://pin.association-aristote.fr/lib/exe/fetch.php/public/presentations/2011/pin20110118_2_aaf_theorie_email_aaf_1701_2011_vd.PDF (consulté le 4 avril 2013).

ASSOCIATION OF RECORDS MANAGERS AND ADMINISTRATORS (ARMA). *Guideline for managing e-mail*. Prairie Village (Kansas) : ARMA International, 2000.

ASSOCIATION POUR L'ECONOMIE NUMERIQUE (ACSEL). *Mille milliard d'e-mails, livre blanc de l'e-mail*. IREPP-ACSEL, 2001, 125 p.

AUCLAIR Valérie. *De nouveaux documents à archiver à l'ère électronique : le cas des courriels*, Université de Montréal, sous la direction de M. Yvon Lemay, mai 2011, 14 p. Consultable en ligne à l'adresse suivante : <http://hdl.handle.net/1866/5075> consulté le 17 avril 2013).

BALLET Philippe, RIETSCH Jean-Marc. *Conserver les courriers électroniques ? Ou comment résoudre la problématique de l'archivage des e-mails*. Paris : FedISA, 2008. consultable en ligne: http://www.fedisa.eu/fedisa2007/fichiers/2008_10_08_15_26_32.PDF (consulté le 12 mai 2013).

BARON, Jason R. *The future of email preservation*. Washington : NARA, 12 mai 2010.

BIA FIGUEIREDO Marie. *Contribution à la compréhension et à la gestion de l'appropriation, le cas de la messagerie électronique*, 2008, thèse de doctorat en sciences de gestion, université Paris-Dauphine.

BIA FIGUEIREDO Marie, KALIKA Michel. *La communication électronique*. Paris : Economica, 2009, 206 p.

BIBLIOTHEQUE ET ARCHIVES NATIONALES DU QUEBEC. *Orientations pour la gestion documentaire des courriels du gouvernement du Québec*, 2009, 20 p. Disponible en ligne: http://www.banq.qc.ca/documents/services/archivistique_ged/Orientations_Gestion_courriels_gouv_Quebec_2009-01-09_VF.PDF (consulté le 27 mars 2013)

BILLSUS Daniel, HILBERT David. *Seamless electronic mail capture with user awareness and consent*. Disponible en ligne: <http://www.google.com/patents?hl=en&lr=&vid=USPATAPP11620850&id=BL2iAAAAEBAJ&oi=fnd&dq=%22email+archiving%22&printsec=abstract#v=onepage&q=%22email%20archiving%22&f=false> (lien vérifié le 20 mars 2013)

BISSONNETTE Natalie. "Gestion des courriels : stratégie, technologies et bonnes pratiques", *Archives*, vol. 44, 2012-2013, n°1, p. 77-113.

BLANCOT Carole, BERTHELOT Vincent, LANSDEHEER Anne de. *Inondé sous les e-mails, résistez !* Paris, Hachette, 2013, 201 p.

BOUDREZ Filip. *Filing and archiving e-mail*. Antwerp, 2006. Disponible en ligne http://www.expertisecentrumdavid.be/docs/filingArchiving_email.PDF (consulté le 12 mars 2013).

BOUKEF CHARKI Nabila, CHARKI M. H. « L'e-mail : un moyen de contrôle ou de responsabilité ? », dans *Systèmes d'Information et Management*, 2008, vol. 13, n°4, p. 29-54.

BOUKEF-CHARKI Nabila, KALIKA Michel, « La théorie du mille-feuille, le rôle du contexte », *Systèmes d'information et management*, 2005, n°11-4, p. 29-54

BRETESCHE Sophie, CORBIERE François de, GEFFROY Bénédicte. « La messagerie électronique, principal métronome des activités de cadre », *La Nouvelle Revue du Travail*, 2012, n°1. Disponible en ligne: <http://nrt.revues.org/262> (Consulté le 27 mai 2013)

BROGAN Mark and VREUDENBURG, S. 'You've Got Mail': *Accountability and End User Attitudes to Email Management*, 4th International Conference on e-Government, RMIT University Melbourne Australia, 2008, pp. 63-69.

BURGESS Anthony, JACKSON Thomas, EDWARDS Janet. « Email training significantly reduces email defects », dans *International Journal of Information management*, 2005, vol. 25, p.71-83.

CENTRE D'ANALYSE STRATEGIQUE, DIRECTION GENERALE DU TRAVAIL. *L'impact des TIC sur les conditions de travail*. Paris : La Documentation française, 2012. 328 p. Disponible en ligne: <http://www.strategie.gouv.fr/content/rapport-l%E2%80%99impact-des-tic-sur-les-conditions-de-travail> (consulté le 12 mai 2013).

CIHUELO Jérôme, SAINTIVE Brigitte. « De l'émiettement du travail intellectuel : déstructuration et dépossession de l'activité de bureau par la messagerie électronique », *Communication & Organisation*, 2000, n° 17. Disponible en ligne: <http://communicationorganisation.revues.org/2365> (consulté le 27 mai 2013)

CLUB DES RESPONSABLES DE POLITIQUE ET PROJETS D'ARCHIVAGE [CR2PA]. *L'archivage des mails ou Les utilisateurs face aux mails qui engagent l'entreprise*. Livre blanc proposé par le groupe de travail "Archivage des mails" du CR2PA. Paris : CR2PA, [2009]. 18 p.

COE-REXECODE. « L'économie numérique et la croissance, poids, impact et enjeux d'un secteur stratégique », document de travail n°24, mai 2011. Disponible en ligne : <http://www.coe-rexecode.fr/public/Analyses-et-previsions/Documents-de-travail/L-economie-numerique-et-la-croissance> (consulté le 27 mai 2013).

COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTES. *Guide pour les employeurs et les salariés*. Paris, 2010. Disponible en ligne: http://www.cnil.fr/fileadmin/documents/Guides_pratiques/CNIL_GuideTravail.PDF (consulté le 27

mai 2013).

CONSEIL D'ANALYSE STRATEGIQUE, « Quel est l'impact des TIC sur les conditions de travail dans la fonction publique ? », *La note d'analyse*, janvier 2013, n° 318, 16 p.

CONSEIL INTERNATIONAL DES ARCHIVES (ICA). *L'archivage, clé d'une bonne gouvernance*, module 17 : *Gérer les courriels*, juillet 2012, 15 p.

COX R. J. "Electronic mail and personal recordkeeping", dans *In Personal Archives and a New Archival Calling: Readings, Reflections and Ruminations*. Duluth (Minnesota) : Litwin Books, chapitre 7, 2008, pp. 201-42.

DEKEYSER Hannelore, DUMORTIER Jos. "L'archivage du courrier électronique au sein de l'administration", dans *Revue du Droit des Technologies de l'Information*, 2004, n° 18, p. 57-65.

DIGITAL PRESERVATION TESTBED. *From digital volatility to digital permanence: Preserving email*. The Hague, avril 2003. Disponible en ligne: <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-email-en.PDF> (consulté le 2 mars 2013).

DIRECTION DES ARCHIVES DE FRANCE. Instruction DITN/RES/2009/007 du 3 juin 2009, consultable en ligne à l'adresse suivante : <http://www.archivesdefrance.culture.gouv.fr/static/2822> (consulté le 15 mai 2013)

DIRECTION INTERMINISTERIELLE DES SYSTEMES D'INFORMATION ET DE COMMUNICATION (DISIC). *Archivage électronique. Un nouveau domaine d'expertise au service de la gouvernance des systèmes d'information*, DISIC, 2012, t. 1, 19 p. Disponible en ligne: [http://references.modernisation.gouv.fr/sites/default/files/DISIC_AE](http://references.modernisation.gouv.fr/sites/default/files/DISIC_AE_Guide_bonnes_pratiques_0.PDF)

[_Guide_bonnes_pratiques_0.PDF](http://references.modernisation.gouv.fr/sites/default/files/DISIC_AE_Guide_bonnes_pratiques_0.PDF)
DISIC. *Archivage électronique. Guide de bonnes pratiques et annexes*, DISIC, 2012, t. 2, p. 18-23. Disponible en ligne: http://references.modernisation.gouv.fr/sites/default/files/DISIC_AE_Annexes_Guide_0.PDF (consulté le 8 mars 2013).

ELSWEILER David, « Understanding the complexities of email behaviour », dans *Information wissenschaft und praxis*, 2012, vol. 63, n°5, p. 314-318.

FAUCHOUX Vincent, DEPREZ Pierre. *Le droit de l'internet. Lois, contrats et usages*. Paris : LITEC, 2008

FERAL-SCHUHL Christiane. *Cyberdroit. Le droit à l'épreuve de l'internet*. Paris : Dalloz, 2010. 6^e édition.

FERRANTE Riccardo, SCHMITZ FUHRIG Lynda. *Digital Preservation : Unsing the Email Account XML Schema*. Washington : Smithsonian Institution, 2007, 8 p.

FLYNN Nancy. *The e-policy handbook : rules and best practices to safely manage your companu's e-mail, blogs, social networking and other electronic communication tools*, New York; Toronto, American management association, 2009, 2^e édition, p. 34-46. Disponible en ligne: http://www.wosco.org/books/Economy/The_e-Policy_Handbook.PDF (consulté le 17 avril 2013).

GAULD Robin, GRAY Andrew, McCOMB Sasha. "How responsiveness is E-gouvernement ? Evidence from Australia and New Zealand", dans *Gouvernement information Quaterly*, 2009, vol. 26, n°1, p. 69-74

GOETHALS Andrea et GOGEL Wendy. "Reshaping the Repository: The Challenge of Email Archiving", dans *7th International Conference on Preservation of Digital Objects (iPRES2010)*, Vienne, Autriche, 2010, p. 71-76. Disponible en ligne: http://publik.tuwien.ac.at/files/PubDat_191968.PDF (consulté le 14 mars 2013).

GRUPO DE TRABAJO DE DOCUMENTOS ELECTRÓNICOS DE LA CONFERENCIA DE ARCHIVEROS DE UNIVERSIDADES ESPAÑOLAS. *Propuesta de Recomendaciones para la Gestión y Conservación del Correo Electrónico en las Universidades Españolas*, Espagne, 2011,

30 p.

HEDSTROM Margaret y WALLACE David. "And the last shall be first: Recordkeeping policies and the NII", dans *Journal of the American Society of Information Science*, vol.50, n°4, 1999, p. 331-339.

HUMAN GENOME ARCHIVES PROJECT. *Review of available tools for managing digital preservation of email*, draft du Wellcome Trust, 2013, 5 p.

HWANG, Yujong. « Understanding moderating effects of collectivist cultural orientation on the knowledge sharing attitude by email », dans *Computers in human behaviours*, 2012, vol. 28, n°6, p. 2169-2174.

ISSA-SALWE A., AHMED M., KABIR M.N., SHARIF L. « Computers mediated communication : mailing list as modern meeting place », dans *Trends in Information management*, 2009, vol. 5, n°2, p. 184-218.

JACKSON Thomas, DAWSON Ray, WILSON Darren. « Reducing the effect of email interruptions on employees », dans *International Journal of Information management*, février 2003, vol. 23, n°1, p.55-65.

JOHN Jeremy Leighton. *Digital forensics and preservation*, DPC Technonology watch report, novembre 2012, 60 p.

KALIKA Michel, BOUTEF CHARKI Nabila, ISAAC Henri. « La théorie du millefeuille et l'usage des TIC dans l'entreprise », *Revue française de gestion*, mars 2007, n° 172, p. 117-129

KIESENDAHL Jana. « Status und Kommunikation : ein Vergleich von Sprechhandlungen », dans *Universitären emails und spechstundengesprächen*, Berlin : E. Schmidt, 2011, 393 p.

KNIGHT J. *InSPECT Project Document - Significant Properties Testing Report : Electronic Mail*. London : JISC/The National Archives/King's College London, 2009. 41 p.

KU Yi-Cheng, CHU Tsai-Hsin, TSENG Chen-Hsiang. « Gratifications for using CMC technologies : a comparison among SNS, IM, and e-mail, computers, dans *Human behaviour*, volume 29, issue 1, janvier 2013, p. 226-234.

KUPPER Beda. "E-Mail Archivierung" dans COUTAZ Gilbert, MEYSTRE-SCHAEREN Nicole, ROTH-LOCHNER Barbara, STEIGMEIER [éd.], *Actualité archivistique suisse. Archivwissenschaft Schweiz aktuell. Travaux du certificat en archivistique et sciences de l'information. Arbeiten aus dem Zertifikat in Archiv- und Informationswissenschaften*. Baden, 2008, pp. 88-117.

LABBE H., MARCOCCIA M. « Communication numérique et continuité des genres : l'exemple du courrier électronique », Version revue de « Tradition épistolaire et médias numériques : du billet au courrier électronique », dans A. Betten & M. Dannerer (eds.), 2005, vol. 2, pp. 281-292.

LE DOUARIN Laurence, « Gérer la vie quotidienne, occuper des postes d'encadrement : les usages « non professionnels » des TIC par les femmes et leur conjoint sur le lieu de travail », *Politiques et Management public*, avril-juin 2011, vol. 28-2, p. 235-252.

LIBRARY of CONGRESS, Extractor tool helps preserves Microsoft Outlook emails and attachments, Library of Congress Digital Preservation, 2010. Disponible en ligne: http://www.digitalpreservation.gov/news/2010/20100924news_article_pedals_email_tool.html (consulté le 14 mars 2013).

LUKESH, Susan, *Email and potential loss to future archives and scholarship or the dog that didn't bark*. Disponible en ligne: [http://hofprints.hofstra.edu/13/01/Lukesh%2C Susan S. \(1999\) E-mail and Potential Loss to Future Archives and Scholarship or The Dog that Didn't Bark. FirstMonday 4\(9\).htm](http://hofprints.hofstra.edu/13/01/Lukesh%2C%20Susan%20S.%20(1999)%20E-mail%20and%20Potential%20Loss%20to%20Future%20Archives%20and%20Scholarship%20or%20The%20Dog%20that%20Didn't%20Bark.%20FirstMonday%204(9).htm) (consulté le 27 mai 2013).

MARKESS INTERNATIONAL. *Prochains enjeux de l'archivage de courriers électroniques*,

avril 2009, 21 p.

MARTINS DA CRUZ José-Marcio. *Spam, classement statistique de messages électroniques : une approche pragmatique*. Paris : Mines Paris-Tech, 2012, 219 p.

McGILL UNIVERSITY ARCHIVES. *Recordkeeping and Email Best Practices*. Disponible en ligne à : http://www.archives.mcgill.ca/dp/assets/PDF_dp_email_best_practices_rev.PDF (consulté le 13 février 2013).

MENNE-HARITZ Angelika. « Managing and archiving administrative records in the digital era- The instrumentality of electronic records for administrative communication », dans *Managing and archiving records in the digital era : changing professional orientations*, sous la direction de Niklaus Bütikofer, Hans Hofman et Seamus Ross, Baden, hier+jetzt, 2006, p.30-38.

MEYER C. “Evolutions in Email Style and Usage”, dans *Science and Technology for Humanity (TIC-STH)*, IEEE Toronto International Conference, 2009, p. 609–612.

MICROSOFT CORPORATION. *Outlook 2013 MAPI Reference*, disponible en ligne à l'adresse suivante : <http://msdn.microsoft.com/en-us/library/cc765775.aspx> (consulté le 10 avril 2013).

MINOR David. *Mail account XML Schema: How Internet messages can be stored as XML*, 2008, disponible en ligne à l'adresse suivante : http://siarchives.si.edu/cerp/David_Minor_CERP_symp.PDF (consulté le 4 avril 2013).

NATIONAL ARCHIVES OF AUSTRALIA. *Managing Electronic Messages as Records*, Disponible en ligne: http://pandora.nla.gov.au/pan/22371/200111050000/www.naa.gov.au/recordkeeping/er/elect_messages/summary.html (consulté le 8 mars 2013).

NATIONAL ARCHIVES OF THE UNITED KINGDOM. *Guidelines on developing a policy for managing emails*, Crown copyright, 2004, disponible en ligne: <http://www.nationalarchives.gov.uk/documents/information-management/managing-emails.PDF> (consulté le 27 mars 2013)

NATIONAL RECORDS AND ARCHIVES ADMINISTRATION. *Bulletin 2013-02 "Guidance on a New Approach to Managing Email Records" draft*. Disponible en ligne : <http://www.archives.gov/records-mgmt/bulletins/2013/2013-02.html> (consulté le 3 septembre 2013).

NOUALI Omar, BLACHE Philippe. « Filtrage automatique de courriels, une approche adaptative et multiniveau », dans *Annales des télécommunications*. Springer, Heidelberg, 2005, volume 60, issue 11-12, p. 1466-1487.

OBSERVATOIRE SUR LA RESPONSABILITE SOCIETALE DES ENTREPRISES (ORSE). *Charte pour un meilleur usage de la messagerie électronique dans les entreprises*. Paris : ORSE, 2011, 30 p. Disponible en ligne: http://www.anact.fr/portal/page/portal/web/publications/NOTINMENU_affichage_document?p_thingIdToShow=22629585 (consulté le 9 avril 2013).

PAPIN Étienne. « La jurisprudence sur le courrier électronique », *Archimag*, décembre 2010-janvier 2011, n° 240, p. 38-39.

PARK Eun G., ZWARISH Nathasha. « Canadian government agencies develop e-mail management policies », dans *International Journal of Information management*, décembre 2008, vol. 28, n°6, p.468-473.

PARTRIDGE Craig. “The Technical Development of Internet Email”, dans *Annals of the History of Computing*, IEEE, 2008, 30 (2), p. 3–29.

PATERSON Giovanna, SPREHE J. Timothy, “Principles challenges facing electronic records management in federal agencies today, dans *Government Information Quarterly*, 2002, vol. 19, n°3, p. 307-315.

PENNOCK Maureen. "Curating E-Mails : A life-cycle approach to the management and preservation of e-mail messages" *DCC Digital Curation Manual*, S.Ross, M.Day (eds.), juillet 2006, disponible en: <http://www.dcc.ac.uk/resource/curation-manual/chapters/curating-e-mails> consulté le 20 février 2013).

PRAS Bernard. « Entreprise et vie privée. Le 'privacy paradox' et comment le dépasser ? », *Revue française de gestion*, 2012, n° 224, p. 87-94

PROM Chris. *Practical E-Records blog*. Disponible en ligne: <http://e-records.chrisprom.com> (consulté le 16 février 2013).

PROM Christopher J. *Preserving email, DPC Technology Watch Report 11-01 December 2011*, 57 p.

REPUBLIQUE ET CANTON DE GENEVE. *Gestion des messages électroniques au sein de l'administration cantonale. Traitement archivistique des messages. Bonnes pratiques concernant l'usage de la messagerie. Logiciel Livelink d'archivage intermédiaire*. Genève : Etat de Genève, 2013. 21 p.

ROBERT Romain. « Archivage des e-mails et protection de la vie privée », dans DEMOULIN Marie [dir.], *L'archivage électronique et le droit*. Bruxelles : Larcier, 2012. 195 p. (collection du CRIDS), p. 55-103

ROMAN H.T. « You've got mail! Understanding how email works, *Tech directions*, 2007, n°66 (7), p. 22-23.

SAINT-LAURENT-KOGAN Anne-France de. « Le travail à l'heure numérique », dans *Cahiers français*, La documentation française, janvier-février 2013, n° 372, p. 24-32.

SAULNIER Carole. « Les courriels : actif informationnel de nos organisations », dans *Archives*, Association des archivistes du Québec, 2005-2006, vol. 37, n°2, p. 119-137.

SMITHSONIAN INSTITUTION ARCHIVES. *The Collaborative Electronic Records Project (CERP)*. Disponible en ligne: <http://siarchives.si.edu/cerp/> (consulté le 15 février 2013).

SOYEZ Sébastien. *Directives pour la gestion et l'archivage numérique des e-mails*, Archives générales du Royaume – Algemeen Rijksarchief, version 1.1, octobre 2009, disponible en ligne : http://arch.arch.be/images/stories/brochure_surveillance/directive_e-mail.PDF (consulté le 10 avril 2013)

SPINDLER Robert P. « Electronic records preservation », dans *Encyclopedia of Library and Information Sciences*, 2010, 3e édition, n°1, p. 1682-1688.

SPURZEM Bob, TOLSON Bill. *Email archiving for dummies*. Hoboken (NJ) : Wiley, 2008, 58 p.

SUMECKI David, CHIPULU Maxwell, OJIAKO Udechukwu. « Email overload : exploring the moderating role of the perception of email as a business critical tool », dans *International Journal of Information management*, octobre 2011, vol. 31, n°5, p.407-414.

SUNLIGHT FOUNDATION. *Sarah's Inbox. A Project of the Sunlight Foundation*, 2011, disponible en ligne: <http://sarahsinbox.com/> (consulté le 20 février 2013).

SYMANTEC CORPORATION. *Enterprise Vault product page*. Disponible en ligne: <http://www.symantec.com/business/enterprise-vault> (consulté le 15 mars 2013).

TAHRI Wadi, FALLEY Bernard. « L'usage de la messagerie électronique : une méta-analyse des travaux francophones sur la période 2000-2008 », *Revue Management et Avenir*, 2010, n°34, p. 183-199.

TASHIRO Hisato, LAU Antonio, MORI Junichiro, FUJII Nobuzumi, KAJIKAWA Yuya. « Email networks and leadership performance », dans *Journal of the American society for information science and technology*, 2012, vol. 63, n°3, p. 600-606.

THE INTERPARES 3 PROJECT, TEAM ITALY. *Guidelines and Recommendations for E-Mail Records Management and Long-Term Preservation*, 2011, version 2.2., 80 p.

TRAN Sébastien. « Quand les TIC réussissent trop bien dans les organisations : le cas du courrier électronique chez les managers », *Management & Avenir*, 2010, 34, p. 200-215

UNIVERSITE D'ABERDEEN. *Guidelines for the Management of University E-Mail*, septembre 2005. Disponible en ligne: http://www.abdn.ac.uk/site-tools/search-results.php?site_search_go=&q=guidelines+for+email+management (consulté le 15 mars 2013).

UNIVERSITE D'EDIMBOURG (UoE), *Managing your emails*, 2005, v. 12. disponible en ligne: http://www.ed.ac.uk/polopoly_fs/1.82267!/fileManager/Managing%20email%20guidance.PDF (consulté le 27 mai 2013).

VAN VLECK Tom. *The history of electronic mail*, 2010. disponible en ligne: <http://www.multicians.org/thvv/mail-history.html> (consulté le 15 mars 2013).

WALLER Aaron David, RAGSDELL Gillian, « The impact of e-mail on work-life balance », dans *New information perspectives*, 2012, vol. 64, n°2, p.154-177.

WEBER R. "The grim reaper: the curse of e-mail", *MIS quarterly*, 2004, vol. 28, n°3, p. 3-14.

WHITTAKER S., SIDNER C. « Email overload: exploring personal information management of email », dans KIESLER S. [éd.], *Culture of the internet*, New Jersey : Lawrence Erlbaum, 1997, p. 227-295.

Annexe 2 : Réglementation et jurisprudence

Réglementation

Réglementation européenne :

- convention européenne des droits de l'homme : article 8 sur le droit au respect de la vie privée ;
- directive 95/46/CE du Parlement européen et du Conseil, du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données ;
- directive 2002/58/CE du Parlement européen et du Conseil du 12 juillet 2002 concernant le traitement des données à caractère personnel et la protection de la vie privée dans le secteur des communications électroniques (directive vie privée et communications électroniques) ;
- projet de règlement européen en matière de protection des données personnelles (en cours de discussion).

Réglementation nationale

code civil :

- article 9 sur le respect de la vie privée ;
- articles 1316-1 à 1316-3 sur la valeur juridique du courrier électronique ;
- article 1316-4 sur la signature électronique.

code du patrimoine

- partie législative et réglementaire, livre II sur les Archives, notamment article L211-1 sur la définition des archives ;

code pénal :

- article 226-1 sur l'atteinte à la vie privée ;
- articles 226-15 et 432-9 sur l'atteinte au secret des correspondances ;
- articles 323-1 à 323-7 relatifs à la fraude informatique.

code des postes et télécommunications

code du travail :

- article **L. 1121-1** sur le principe du respect par l'employeur des libertés individuelle et collective ;
- articles **L. 1221-9 et 1222-4** dispose qu'aucune information ne peut-être collectée par un dispositif qui n'a pas été porté à la connaissance du salarié ;
- article **L 2323-32** impose l'information et la consultation du comité d'entreprise avant la mise en œuvre dans l'entreprise de moyens de contrôle de l'activité du salarié ;
- article **L 2142-6** sur la diffusion syndicale dans l'entreprise, qui précise que l'utilisation de la messagerie électronique de l'entreprise pour la diffusion des tracts électroniques doit faire l'objet d'un accord-cadre avec l'entreprise et ne doit pas nuire au bon fonctionnement du réseau informatique ni entraver l'accomplissement du travail.

Autres textes :

- loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, modifiée par la loi n° 2004-801 du 6 août 2004 ;
- décret n° 2001-272 du 30 mars 2001 pris pour l'application de l'article 1316-4 du code civil et relatif à la signature électronique.

Jurisprudence

Arrêts de la Cour de cassation :

Courrier électronique et son utilisation par le salarié

Courriers personnel et professionnel :

- 2 octobre 2001 (arrêt Nikon). L'arrêt énonce pour la première fois que le salarié a droit, même au temps et au lieu de travail, au respect de l'intimité de sa vie privée ;
- 18 octobre 2006 sur l'identification du caractère personnel des échanges par le salarié ;
- 30 mai 2007 sur les correspondances échangées sur le lieu de travail sans élément professionnel ;
- 6 juin 2007. L'employeur peut prendre connaissance des courriels professionnels des salariés, mais pas des courriels identifiés comme personnels ;
- 12 mai 2010 sur le licenciement d'un salarié ayant manipulé un autre salarié par l'intermédiaire de faux courriels ;
- 19 mai 2010 sur le bien-fondé d'une mise à pied prononcée à l'encontre d'une salariée, déléguée syndicale, qui avait utilisé la messagerie électronique de l'entreprise pour la distribution de tracts syndicaux au-delà du quota ;
- 8 octobre 2011 sur la présomption du caractère professionnel des courriels échangés grâce au matériel de l'entreprise ;
- 10 janvier 2012 sur l'envoi de courrier électronique par un délégué syndical non reconnu comme obligatoirement relevant de la catégorie « tract électronique » ;
- 4 avril 2012 sur la protection des salariés devant disposer sur leur lieu de travail d'un matériel excluant l'interception de leurs communications téléphoniques et l'identification de leurs correspondants (violation de la loi CNIL de 1978). Cette décision pourrait être étendue à la messagerie électronique et aux logs de connexions internet depuis le poste du salarié ;
- 12 février 2013 (pourvoi n° 11-28649) : Une clé USB, dès lors qu'elle est connectée à un outil informatique mis à la disposition du salarié par l'employeur pour l'exécution du contrat de travail, est présumée utilisée à des fins professionnelles. En conséquence, les dossiers et fichiers non identifiés comme personnels qu'elle contient, sont présumés avoir un caractère professionnel de sorte que l'employeur peut y avoir accès hors la présence du salarié ;
- 19 juin 2013 (pourvoi n° 12-12138) : Des courriels et fichiers intégrés dans le disque dur de l'ordinateur mis à disposition du salarié par l'employeur ne sont pas identifiés comme personnels du seul fait qu'ils émanent initialement de la messagerie électronique personnelle du salarié.

Courrier électronique et liberté d'expression :

- 10 novembre 2009 sur la liberté d'expression « sans intention de nuire » des salariés par courriel ;

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

- 15 décembre 2009 sur la limite de la liberté d'expression d'un salarié par l'emploi de termes « excessifs et injurieux ».

Courrier électronique et droits de l'employeur

Les droits de l'employeur :

- 14 mars 2000 sur le principe de loyauté de l'employeur, d'avertir les salariés d'un possible contrôle des courriels professionnels ;
- 19 mai 2004 sur le contrôle de l'employeur des activités de ses salariés pendant le temps de travail ;
- 12 octobre 2004 sur les limites aux modalités d'accès par l'employeur aux fichiers informatiques ;
- 5 juillet 2011 sur l'utilisation par l'employeur de courriels qui n'ont pas été identifiés comme « personnels » par le salarié ;
- 26 juin 2012 sur la stricte application des garanties formelles inscrites dans le règlement intérieur de la charte informatique de l'entreprise.

Les obligations de l'employeur :

- 28 décembre 2001 sur la recevabilité de réclamation par courrier électronique ;
- 15 octobre 2003 sur le principe de laïcité des fonctionnaires appliqué au courrier électronique ;
- 2 juin 2004 sur l'utilisation de la messagerie professionnelle pour diffuser des propos antisémites entraînant un licenciement pour faute grave ;
- 7 avril 2006 sur la prise en compte d'un courriel pour sanctionner un fonctionnaire ;
- 23 mai 2007 sur la possibilité pour l'employeur d'obtenir une ordonnance autorisant un huissier à accéder aux données contenues dans l'ordinateur d'un de ses salariés s'il a des motifs légitimes de suspecter une concurrence déloyale ;
- 17 juin 2009 sur la procédure d'information et de consultation préalable du Comité d'entreprise avant la mise en œuvre de moyens ou de techniques permettant le contrôle de l'activité des salariés ;
- 14 avril 2010 sur le non licenciement d'un salarié possédant des documents pornographiques sur le disque dur de son ordinateur professionnel.

Arrêts de cours d'appel :

- avril 2003 : la cour d'appel de Paris a condamné un salarié pour atteinte portée à l'image de marque et à la réputation de son entreprise ;
- juillet 2003 : la cour d'appel de Bordeaux a reconnu aux courriels leur caractère personnel, même s'ils étaient émis depuis l'adresse générique de l'entreprise (et non une adresse individuelle) ;
- 14 janvier 2010 : la cour d'appel de Rennes estime que les courriels envoyés par un fonctionnaire depuis sa messagerie de travail sont « *a priori* » de nature professionnelle.

Annexe 3 : Études de cas

Ministère des Affaires étrangères

Catégorie de paramètre	Possibilités de paramétrage	Impact sur l'usage
Mode de gouvernance du système	<ul style="list-style-type: none"> - une adresse par individu (nom.prenom@diplomatie.gouv.fr). - possibilité de créer des adresses fonctionnelles. - un outil commun à l'administration centrale et aux postes à l'étranger. 	
Degré d'ouverture du système	<ul style="list-style-type: none"> - ouverture sur le réseau internet. - transfert automatique des courriels vers l'extérieur interdit. - mise à disposition aux agents de token pour se connecter au SI à distance et accéder ainsi à la messagerie et à l'intranet. - une cinquantaine de téléphones de type Hermés (agréé ANSSI) permettent de recevoir et d'envoyer sur la messagerie interne. Matériel non ergonomique qui n'est pas apprécié. 	<ul style="list-style-type: none"> - Utilisation de la messagerie impossible sur tablette et smartphone personnel. - Seuls les personnes placées le plus haut hiérarchiquement ont accès aux messageries en mobilité.
Environnement technique	<ul style="list-style-type: none"> - Microsoft EXCHANGE à la Centrale et Kerio Serveur à l'étranger. - absence de messagerie instantanée - absence de solution d'ERMS et d'outil de travail collaboratif 	Les agents du ministère utilisent la messagerie comme outil de discussion, comme outil de travail collaboratif et comme outil de mémoire institutionnelle.
Protocole d'accès	- protocole Exchange + IMAP	
Choix et contraintes en matière de stockage	- stockage bridé sur la boîte de réception à 80 Mo pour les utilisateurs et à 500 Mo pour les VIP.	Les utilisateurs créent des dossiers locaux stockés que les groupes bureautiques.
Choix et contraintes en matière d'accès	<ul style="list-style-type: none"> - <i>webmail</i> « OWA » Outlook Web Access pour les agents de la Centrale et un <i>webmail</i> standard pour les agents à l'étranger ; - client de messagerie Outlook 2003 ou 2010 installé sur le poste de travail 	
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : classement et capacité de créer des dossiers locaux	<ul style="list-style-type: none"> - stockage des courriels sur des PST ; - aucune obligation de respecter un plan de classement pré-établi ; - aucun classement automatique. 	Chaque utilisateur structure sa messagerie selon sa logique propre
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : impression et export	<ul style="list-style-type: none"> - format d'export du client Outlook : eml, pst ; - possibilité d'imprimer en PDF <i>via</i> PDFCreator. 	
Accompagnement des utilisateurs	<ul style="list-style-type: none"> - formations et supports de communication ; - <i>a priori</i> aucune instruction spécifique au MAE sur l'utilisation des messageries. 	

Exemple de la messagerie d'un agent du ministère des Affaires étrangères

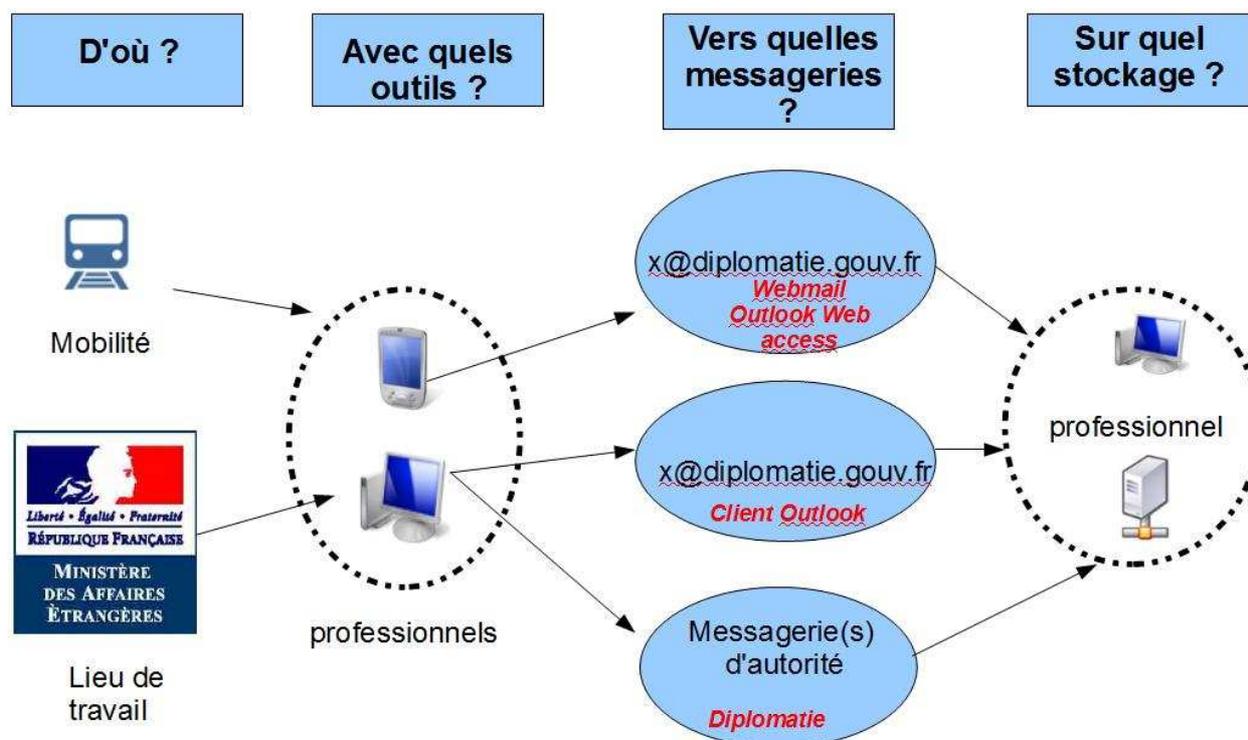


Fig. 9 : modélisation des possibilités d'utilisation des messageries par un agent du ministère des Affaires étrangères

Ministère de la Culture et de la Communication

Le paramétrage technique de la messagerie utilisée au ministère de la Culture et de la Communication est le suivant :

Catégorie de paramètre	Possibilités de paramétrage	Impact sur l'usage
Mode de gouvernance du système	<ul style="list-style-type: none"> - une adresse par individu (nom.prenom@culture.gouv.fr). - possibilité de créer des adresses fonctionnelles. - un outil commun à l'administration centrale, aux services déconcentrés et aux services à compétence nationale 	
Degré d'ouverture du système	<ul style="list-style-type: none"> - ouverture sur le réseau internet 	Les agents ont la possibilité de transférer des informations sur leur messagerie personnelle.
Environnement technique	<ul style="list-style-type: none"> - absence de messagerie d'autorité ; - absence de messagerie instantanée ; - absence de solution d'ERMS et d'outil de travail 	Les agents du ministère utilisent la messagerie comme outil de discussion, comme outil de travail collaboratif et comme outil de mémoire

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

	collaboratif (à la seule exception des groupes de travail accessibles depuis l'intranet Sémaphore aux seuls utilisateurs référencés dans l'annuaire du ministère)	institutionnelle.
Protocole d'accès	- protocole IMAP	Les agents qui le souhaitent peuvent consulter leur messagerie depuis leurs outils personnels (smartphone, ordinateur, tablette, etc.) ou <i>via</i> leurs outils professionnels.
Choix et contraintes en matière de stockage	- stockage bridé sur le <i>webmail</i> à 500 Mo	Les agents créent des dossiers locaux stockés sur leur disque dur
Choix et contraintes en matière d'accès	- <i>webmail</i> « Telemac » accessible <i>via</i> une connexion sécurisée ; - client de messagerie Thunderbird installé sur le poste de travail (hors cabinet ministériel) ; - client de messagerie Outlook installé sur le poste de travail (cabinet ministériel).	Les utilisateurs peuvent accéder à leurs messages à distance, sur leurs équipements professionnels comme personnels. Ils peuvent en conséquence stocker des messages sur des espaces de stockage professionnels comme personnels.
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : classement et capacité de créer des dossiers locaux	- aucune obligation de respecter un plan de classement pré-établi ; - aucun classement automatique	Chaque agent structure sa messagerie selon sa logique propre.
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : impression et export	- format d'export du client Thunderbird : eml, mbox ; - format d'export du client Outlook : pst, msg ; - absence d'interfaçage avec une imprimante virtuelle.	
Accompagnement des utilisateurs	- formations techniques dispensées par la SDSI dans le plan de formation ; - diffusion de préconisations pour une bonne gestion de la messagerie par la Mission des archives sur l'intranet du ministère, Sémaphore ¹⁸⁶ ; - diffusion par les Archives nationales, service à compétence nationale du ministère, d'un vademecum propre à destination de leurs agents et disponible à la fois sur l'intranet du ministère et sur le site internet de l'institution (rubrique « Publications professionnelles » ¹⁸⁷), avec un accompagnement sous forme de formations.	

¹⁸⁶ <http://semaphore.culture.gouv.fr/web/mes-dossiers-au-quotidien/gerer-ma-messagerie> (consulté le 14/05/2013).

¹⁸⁷ Disponible en ligne : <http://www.archivesnationales.culture.gouv.fr/chan/chan/PDF/livret-vade-courriels.PDF> (consulté le 14/05/2013).

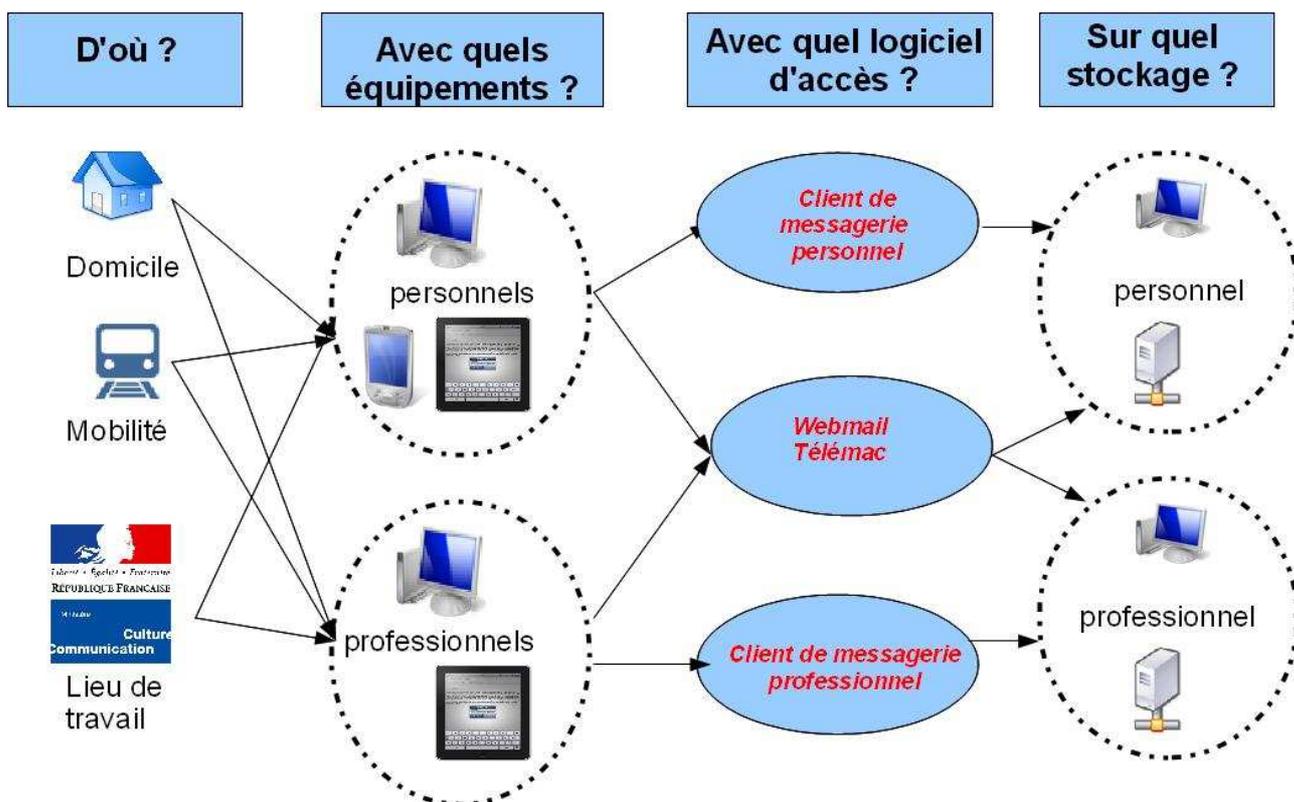


Fig. 10 : modélisation des possibilités d'utilisation des messageries par un agent du ministère de la Culture et de la Communication

Ministère de la Défense

Catégorie de paramètre	Possibilités de paramétrage	Impact sur l'usage
Mode de gouvernance du système	<ul style="list-style-type: none"> - plusieurs adresses possibles par individu : <ul style="list-style-type: none"> ^ individus arrivés avant la migration sur le socle technique commun : au minimum une adresse intranet (prenom.nom@direction.gouv.fr) et une adresse internet (prenom.nom@defense.gouv.fr) ; ^ individus arrivés après la migration sur le socle technique commun : une seule adresse (prenom.nom@intradef.gouv.fr) 	
Degré d'ouverture du système	- 2 réseaux séparés : réseau internet / réseau intranet (accès par 2 postes de travail dissociés) ;	Les agents ont la possibilité de transférer des informations sur leur messagerie personnelle. Cette pratique a tendance à croître faute d'accès possible à la messagerie intradef depuis le réseau internet (impossibilité de consulter les liens hypertexte, impossibilité d'accéder aux messages depuis un équipement mobile ou un poste non relié au réseau intranet).

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

Environnement technique	<ul style="list-style-type: none"> - existence de messageries d'autorité ; - existence d'outils de travail collaboratifs (agoras Lotus Notes pour la DGA, espaces Sharepoint pour EMA) ; - diffusion dans le nouveau socle technique commun d'outils de messagerie instantanée, d'une webcam (en fonction du profil de l'utilisateur). 	<p>Les agents ont à leur disposition sur leur poste de travail intranet un ensemble d'outils de gestion de l'information.</p> <p>Aucun outil n'est disponible depuis un poste connecté à l'internet (du moins pour un utilisateur SGA).</p>
Protocole d'accès	- protocole IMAP	Les agents qui le souhaitent peuvent consulter leur messagerie « defense » depuis leurs outils personnels (smartphone, ordinateur, tablette, etc.) ou <i>via</i> leurs outils professionnels.
Choix et contraintes en matière de stockage	- <i>webmail</i> « intradef » bridée à 200 Mo ;	Les agents créent des dossiers locaux stockés sur leur disque dur.
Choix et contraintes en matière d'accès	<ul style="list-style-type: none"> - <i>webmail</i> ; - client de messagerie installé sur le poste de l'utilisateur : <ul style="list-style-type: none"> ▲ client Lotus Note pour la DGA avant la migration sur le socle technique commun ; ▲ client Thunderbird pour le SGA avant la migration sur le socle technique commun ; ▲ client Outlook après la migration sur le socle technique commun. 	<p>Les agents peuvent accéder à leurs messages « defense » à distance, sur leurs équipements professionnels comme personnels.</p> <p>Ils peuvent en conséquence stocker des messages sur des espaces de stockage professionnels comme personnels.</p> <p>Ce n'est pas le cas pour les agents qui ne disposent que d'adresses « directions » ou « intradef » et qui ne peuvent accéder à leurs messages que depuis un équipement professionnel.</p>
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : classement et capacité de créer des dossiers locaux	<ul style="list-style-type: none"> - aucune obligation de respecter un plan de classement pré-établi ; - aucun classement automatique 	Chaque agent structure sa messagerie selon sa logique propre.
Fonctionnalités du logiciel retenu (<i>webmail</i> et client) : impression et export	<ul style="list-style-type: none"> - format d'export du client Thunderbird : eml, mbox ; - format d'export du client Lotus Note : .nsf, mbox, eml ; - format d'export du client Outlook : pst, msg ; - interfaçage possible avec une imprimante virtuelle (PdfCreator) ; - client de messagerie programmé pour forcer les utilisateurs à constituer des « archives » automatiquement. 	
Accompagnement des utilisateurs	- documents divers disponibles sur l'intranet du ministère et formations techniques assurées par le CFD.	

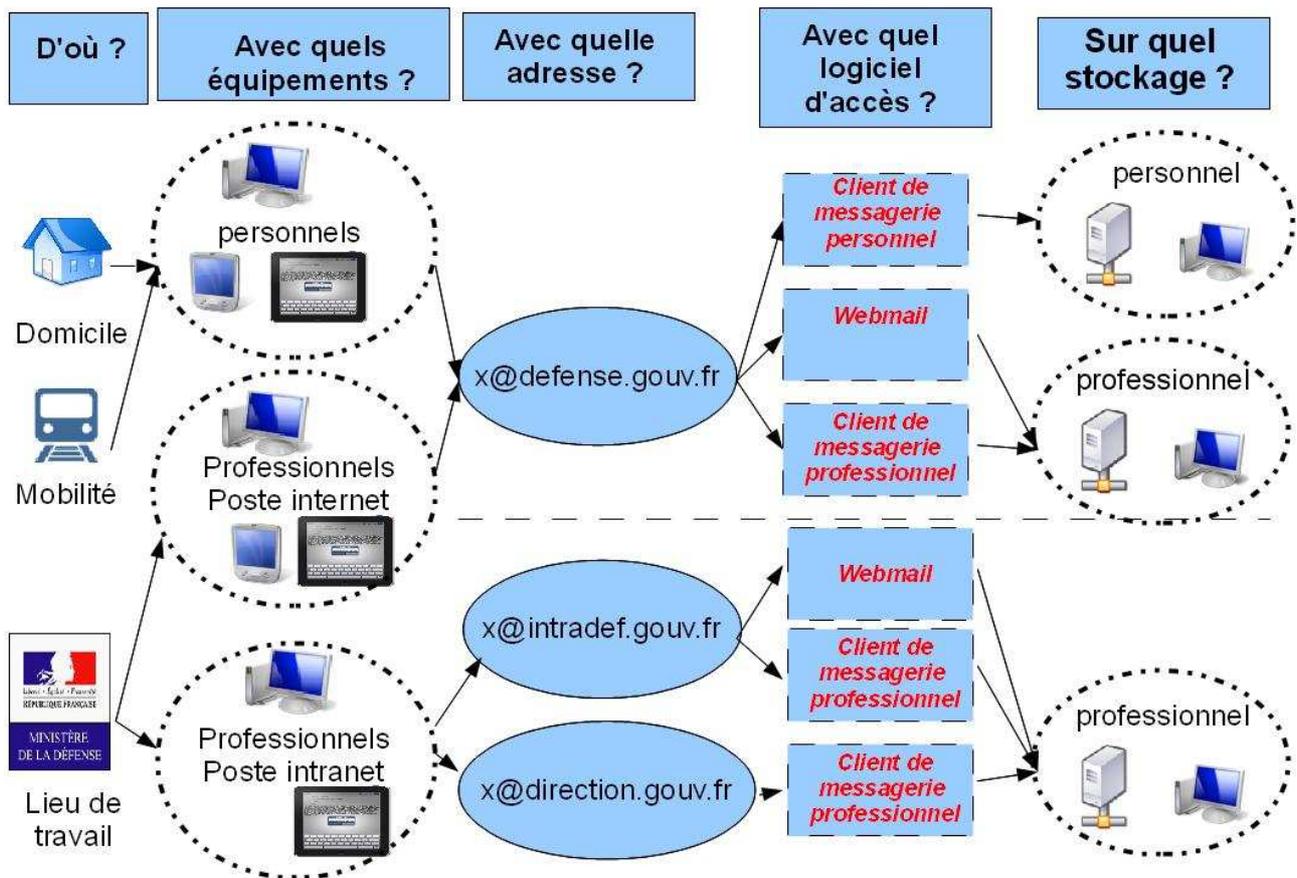


Fig. 11 : modélisation des possibilités d'utilisation des messageries par un agent du ministère de la Défense

Annexe 4 : Les outils utilisés dans le cadre des tests

Outil 1 testé : Aid4mail

Nom de l'outil	Aid4mail
Source URL	http://www.commentcamarche.net/download/telecharger-34058579-aid4mail http://www.aid4mail.com
Formats traités	.cnm, .dbx, .eml, .emlx, .idb, .idx, .mbs, .mbx, .mht, .msg, .nws, .pmi, .pmm, .pmx, .pst, .tbb
Technologie	
Operating system	Windows® 8 / 7 / Vista / XP / 2000, Server 2008 / 2003 ou Mac OS X / Linux avec Wine .
Dépendances	Wine pour la version Mac
Licence d'utilisation	Payant (Fookes Software Ltd)
Catégorie	Conversion
Description de l'outil	Aid4Mail est capable de convertir de nombreux formats de messages dans un large nombre de formats ouverts ou propriétaires, avec un haut degré de fidélité. Aid4Mail peut également se connecter à des serveurs IMAP ou MAPI (comme GMail et Microsoft Exchange) pour moissonner directement les messages. Il comprend des fonctionnalités de filtrage pour exclure ou inclure des messages correspondant à certains critères, un langage de scripts qui permet de configurer des formats d'export et permet la conversion au format PDF/a. C'est un programme stable, qui permet de convertir rapidement des comptes, même volumineux, et il est disponible avec une licence "Archiviste" conçue pour les institutions patrimoniales qui cherchent à conserver des comptes pour des objectifs patrimoniaux ¹⁸⁸ .
Dernière mise à jour	Version 3.0 (juin 2013)
Commentaires	Testé dans le cadre du projet CARCANET. Version gratuite de test avec des fonctionnalités réduites.

Outil 2 testé : Apache POI

Nom de l'outil	Apache POI
Source URL	http://poi.apache.org/
Formats traités	Microsoft : OOXML, XLSX, DOCX, PPTX, OLE2, XLS, DOC, PPT, PST, Visio, Publisher, Winmail.dat (Outlook, Exchange)
Technologie	Java 1.5 minimum
Operating system	tous
Dépendances	Java 1.5 minimum et autres modules Java (voire le site)
Licence d'utilisation	Apache 2.0

¹⁸⁸PROM Christopher J. *Preserving Email (...)*, p. 26.

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

Catégorie	Outil de manipulation et d'extraction de contenu
Description de l'outil	Outil permettant la manipulation et l'extraction de contenu de différents formats Microsoft
Dernière mise à jour	3.9 du 3 décembre 2012
Commentaires	Nécessite de la programmation pour être utilisé puisqu'il s'agit d'une bibliothèque et non d'un produit autonome

Outil 3 testé : CERP Email parser

Nom de l'outil	CERP Email Parser
Source URL	http://siarchives.si.edu/cerp/parserdownload.htm
Formats traités	mbox
Technologie	VM open source Smalltalk Squeak
Operating system	Windows, Mac
Dépendances	
Licence d'utilisation	MIT License pour le parser. La documentation d'utilisation relève de la MIT License de la licence Creative Commons Attribution Non-commercial Share Alike 3.0 .
Catégorie	Conversion
Description de l'outil	Application web qui tourne sur des machines virtuelles <i>open source</i> , CERP Email Parser permet de transformer un ou plusieurs fichiers au format mbox dans un fichier au format XML qui contient le contenu d'un compte entier, conforme aux exigences du format XML de compte de messagerie (E-mail Account Schema). Le logiciel a été développé conjointement par la Smithsonian et les Archives de Caroline du Nord. Il peut être téléchargé sur le site du projet (http://siarchives.si.edu/cerp/). L'application fournit plusieurs options pour la manipulation et l'encodage des pièces jointes. Des développements complémentaires permettront de rendre l'outil plus facile à utiliser et doivent éliminer les goulots d'étranglement dans la conversion qui affectent la performance de l'outil lors de la conversion de comptes volumineux ¹⁸⁹ .
Dernière mise à jour	Décembre 2008
Commentaires	Uniquement testé par le Smithsonian Institution Archives avec Windows XP. Cité dans le rapport <i>DPC Watch Report</i> . En cas d'utilisation et de test, envoyer un email à IT Archivist Ricc Ferrante (ferranter [at] si.edu)

Outil 4 testé : EmailChemmy

Nom de l'outil	EmailChemmy
Source URL	http://www.weirdkid.com/products/emailchemmy/#version
Formats traités	AOL for Windows ("PFC" files), Apple Mail, Claris EMailer for Macintosh, CompuServe Classic for Macintosh (MacCIM), CompuServe for Windows (WinCIM and 2000), Entourage (Database, .RGE Archives and cache files), Eudora, Mozilla, Mulberry, Musashi,

¹⁸⁹*Ibid.*, p. 27.

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

	Neoplanet, Netscape, Opera, Outlook for Mac (8.x, 2001, and 2011 OLM files), Outlook for Windows (MSG, PST and OST files), Outlook Express for Macintosh, Outlook Express for Windows, Outlook Express for UNIX/Solaris, Outspring, PowerTalk/AOCE for Macintosh, QuickMail Pro for Macintosh, QuickMail Pro for Windows, Thunderbird, Windows Live Mail, Windows Mail, Yahoo! Mail, tous les courriels au format mbox ou UNIX
Technologie	Java
Operating system	tous
Dépendances	
Licence d'utilisation	End User licence
Catégorie	Conversion, migration de clients de messagerie dont Outlook et tous ceux qui utilisent leur propre serveur IMAP. Import des messageries dans les applications Google.
Description de l'outil	Application Java propriétaire, EmailChemy permet de convertir des messages depuis de nombreux formats propriétaires et ouverts, enregistrés sur disque dur dans des formats cibles ouverts comme eml et mbox. Il inclut également un serveur IMAP et peut migrer tout message vers un serveur IMAP. Des versions Windows, Mac, Linux/Solaris/Unix sont disponibles (http://www.weirdkid.com/products/emailchemy/) ¹⁹⁰
Dernière mise à jour	Version 11.3.5. du XXX
Commentaires	2 Go de RAM nécessaire à son bon fonctionnement. Conversion vers un format reconnu par la norme RFC-2822. Coût : Migration version=> prix sur demande pour les institutions (< à 269 \$)

Outil 5 testé : Java LibPST

Nom de l'outil	Java LibPST
Source URL	https://github.com/rjohnsondev/java-libpst
Formats traités	pst (Outlook)
Technologie	Java
Operating system	tous
Dépendances	Modules Java
Licence d'utilisation	Apache 2.0 et LGPL
Catégorie	Extraction de contenus
Description de l'outil	Outil d'extraction de contenu au format pst
Dernière mise à jour	0.8 beta 18 décembre 2012
Commentaires	Nécessite de la programmation pour être utilisé puisqu'il s'agit d'une bibliothèque et non d'un produit autonome

Outil 6 testé : Java Mail

¹⁹⁰*Ibid.*, p. 27.

Programme VITAM. L'archivage des messageries électroniques / Preuve de concept

Nom de l'outil	Java Mail
Source URL	http://www.oracle.com/technetwork/java/javamail/index.html et https://java.net/projects/javamail/pages/Home
Formats traités	EML
Technologie	Java 1.6 conseillé
Operating system	tous
Dépendances	Java 1.6 ou sinon d'autres dépendances (JAF)
Licence d'utilisation	CDDL 1.1 et GPL V2 (sauf Classpath)
Catégorie	Outil de manipulation, création et extraction de courriel
Description de l'outil	Permet la création, l'extraction ou la manipulation de courriel au format eml, y compris l'envoi
Dernière mise à jour	1.5.0 au 8 avril 2013
Commentaires	Nécessite de la programmation pour être utilisé puisqu'il s'agit d'une bibliothèque et non d'un produit autonome Pour gérer les fichiers au format mbox, il est nécessaire d'extraire les fichiers au format eml programmatiquement avant de les passer à Java Mail. Pour cela, il est conseillé d'utiliser le principe d'implémentation proposé par https://github.com/nicbet/MailboxMiner et en particulier https://github.com/nicbet/MailboxMiner/blob/master/MailboxMiner2/src/ca/queensu/cs/sail/mailboxmina2/common/MailboxParser.java

Outil 7 testé : XENA (XML Electronic Normalising of Archives)

Nom de l'outil	XENA (XML Electronic Normalising of Archives)
Source URL	http://xena.sourceforge.net/
Formats traités	pst, Mbox, .trim
Technologie	Java
Operating system	tous
Dépendances	aucune
Licence d'utilisation	GNU (general public license)
Catégorie	Conversion, description
Description de l'outil	Outil développé par les Archives nationales d'Australie, Xena utilise la bibliothèque readpst pour convertir des dossiers personnels Outlook au format mbox. Xena est disponible à l'adresse suivante : http://xena.sourceforge.net/191 Outil en base Java développé par les Archives nationales d'Australie pour convertir les courriels en XML dans un but de pérennisation.
Dernière mise à jour	Version 6.0.1 du 14 juin 2012
Commentaires	Utilise ReadPST pour les fichiers produits sous Outlook.

191 *Ibid.*, p. 28.

Annexe 5 : Résultat des tests d'extraction des métadonnées

Champ	RFC 5322 Communs aux 2 formats	EML spécifique Notam. MIME RFC 2045 à 2049	PST spécifique	Description	MBOX	EML	PST	MSG
					Java Mailx	Java Mailx	Java LibPST	Apache POI
Métadonnées extraites								
FolderName			X-Folder :	Nom du dossier	ok (Nom du fichier et chemin)		ok	partiel (dernier niveau uniquement)
From:	From:			émetteur du message	ok	ok	ok	ok
Sender:	Sender:			émetteur du message (parfois différent pour des raisons de secrétariat)	ok	ok		
Reply-To:	Reply-To:			émetteur du message (parfois différent pour des raisons de liste de diffusion)	ok	ok		ok
To:	To:			destinataires principaux	ok	ok	ok	ok
Cc:	Cc:			destinataires en copie	ok	ok	ok	ok
Bcc:	Bcc:			destinataires en copie cachée	ok	ok	ok	ok
display-name	display-name			Pour une adresse courriel : nom affichable	ok	ok	ok	ok
addr-spec	addr-spec			Pour une adresse courriel : adresse effective	ok	ok	ok	ok
Subject:	Subject:			Sujet du message	ok	ok	ok	ok
sentDate	Date:			Date d'émission	ok	ok	ok	ok
receivedDate		via Received:	X-OriginalArrivalTime	Date de réception	ok	ok	ok	ok
trace (Received)	Received:			Traces de la réception du message (serveurs SMTP)	ok	ok	ok	ok
emailSize		natif	X-SDOC: et natif	Taille du courriel	ok	ok	ok	ok
Message-ID:	Message-ID:			Message-ID (identifiant unique) du courriel	ok	ok	ok	ok
In-Reply-To:	In-Reply-To:			Chaîne de réponse (fil) du courriel	ok	ok	ok	ok
Importance:	Importance:			Niveau d'importance du courriel			ok	ok
Priority:	Priority:		Ou X-Priority:	Priorité du courriel (similaire à Importance)	ok	ok	ok	ok
Sensitivity:	Sensitivity:			Sensibilité du courriel (dans un contexte d'entreprise)	ok	ok	ok	ok
hasAttachment		natif	natif	Ce courriel a-t-il des pièces jointes	ok	ok	ok	ok
creationTime (attachment)		modification-date=	natif	Date de création d'une pièce jointe	ok	ok	ok	KO
filename (attachment)		filename=	natif	Nom de la pièce jointe	ok	ok	ok	ok
size (attachment)		natif	natif	Taille de la pièce jointe	ok	ok	ok	ok
attachment as email		natif	natif	Capacité à extraire les informations d'un email en PJ	ok	ok	ok	ok
extraction des corps de texte		Natif (2 formats via Content-Type : TXT et HTML)	Natif (3 formats : TXT, HTML, RTF)	Capacité à extraire le corps du message (pour pérennisation)	ok	ok	ok (TXT, HTML ou RTF)	ok (TXT, HTML ou RTF)
extraction des PJ		natif	natif	Capacité à extraire la PJ (pour pérennisation)	ok	ok	ok	ok
Programmation								
extraction de keywords				Capacité à extraire les mots clés	ok	ok	ok	ok
extraction des fils de messages				Capacité à reconstituer les fils de messages	ok	ok	ok	ok
identification de l'objet (Droid)				Capacité à identifier l'objet primaire via Droid	NON mais un test via Mailx est fait	ok	ok	ok
identification des PJ (Droid)				Capacité à identifier l'objet PJ via Droid	ok	ok	ok	ok

Annexe 6 : Glossaire

Les définitions suivantes ont été établies à partir des glossaires :

- des normes ISO 14721 (OAIS), ISO 16575-2 (ICA-Req),
- des publications relatives à l'archivage électronique suivantes :

ARCHIVES NATIONALES. La gestion et l'archivage des courriels, manuel pratique, version 2, août 2012, p. 51-54.
CERP, Email guidelines, 2010, p. 20-22

CINES : site internet, rubrique sur les métadonnées de pérennisation, www.cines.fr/spip.php?rubrique231

DIGITAL PRESERVATION TESTBED. From digital volatility to digital permanence: Preserving email. The Hague, avril 2003, p. 22-23.

DIRECTION INTERMINISTERIELLE DES SYSTEMES D'INFORMATION ET DE COMMUNICATION (DISIC). Archivage électronique. Un nouveau domaine d'expertise au service de la gouvernance des systèmes d'information, DISIC, 2012, t. 1, p. 82-84. Disponible en ligne: http://referencemodernisation.gouv.fr/sites/default/files/DISIC_AE_Guide_bonnes_pratiques_0.PDF

KNIGHT J. InSPECT Project Document - Significant Properties Testing Report : Electronic Mail. London : JISC/The National Archives/King's College London, 2009. 41 p.

PROM Christopher J. Preserving email, DPC Technology Watch Report 11-01 December 2011, p.38-40.

SMITHSONIAN INSTITUTION ARCHIVES. The Collaborative Electronic Records Project (CERP). Email guidelines, 2010, p. 20-22.. Disponible en ligne: <http://siarchives.si.edu/cerp/>.

- le vocabulaire spécifique emprunté au modèle conceptuel fonctionnel et technique de la solution VITAM est indiqué par « Programme VITAM ».

Adresse électronique : chaîne de caractères constituée de l'identifiant de l'utilisateur et de celui du gestionnaire d'un service de messagerie électronique séparés par le caractère @, identifiant de manière unique cet utilisateur ;

Archival information Package (AIP) : paquet d'information archivé. Cet acronyme désigne, au sein de la norme OAIS, les archives gérées et conservées au sein du système d'archivage électronique ;

Bibliothèque open source : ensemble de bibliothèques de composants logiciels *open source* permettant de répondre, dans le cas des messageries, aux besoins fonctionnels d'extraction des corps, des pièces jointes et des métadonnées descriptives et techniques des courriels ;

Big Data : énorme volume de données très variables qui ne peuvent être gérées par les technologies traditionnelles (notamment SQL) de l'information. Pour y répondre, des nouvelles technologies ont été créées : les bases de données NoSQL, les infrastructures de stockage dite *Cloud computing* par exemple ;

Client de messagerie : logiciel installé localement sur un ordinateur qui sert à envoyer et lire des courriers électroniques. Exemples : Thunderbird, Outlook, Lotus Note ;

Cloud computing ou informatique en nuage : forme particulière de gérance de l'informatique, dans laquelle l'emplacement et le fonctionnement du nuage ne sont pas portés à la connaissance du client et qui permet d'accéder *via* un réseau de télécommunications, à la demande et en libre service, à des ressources partagées configurables (JORF, n°0129 du 6 juin 2010 page 10453) ;

Compte de messagerie : caractérisé par une adresse de messagerie unique, il regroupe l'intégralité des messages reçus et envoyés au moyen de cette adresse ;

Corps (*body*): contenu du message non structuré situé après l'en-tête. Il peut être complété par des pièces jointes attachées au message ;

Décommissionnement : retrait de service d'une application qui peut se traduire soit par un archivage de tout ou partie des données et/ou par une purge de tout ou partie des données. Dans le cadre de la gestion du cycle de vie des données, il est nécessaire de mettre en place des procédures de décommissionnement des applications ;

Dissemination information package (DIP) : paquet d'information diffusé. Il désigne les versions de diffusion des archives conservées dans le système d'archivage électronique ;

Digital Record Object Identification (DROID) : outil développé par les Archives nationales du Royaume-Uni (The National Archives, TNA). Il permet la reconnaissance des formats informatiques en s'appuyant sur le référentiel des formats PRONOM, lui aussi administré par TNA ;

En-tête (*header*) : elle comprend les informations relatives à la rédaction, à l'expédition, à la transmission et à la gestion des messages : identité de l'expéditeur ; identité des destinataires (destinataires principaux, destinataires pour information, destinataires cachés) ; objet ; dates (envoi, réception par les différents serveurs intermédiaires, réception) ; niveau de priorité et de sensibilité ; accusé de réception ; suivi (indexation relative à l'action à réaliser par le destinataire) ;

Encodage : manière de représenter les caractères (lettres, chiffres, symboles) dans un système informatique (Wikipédia) ;

Electronic records management system (ERMS) : logiciel de gestion électronique de documents qui comporte des fonctionnalités de gestion du cycle de vie des documents (CERP) ;

Fil de discussion : suite de messages consécutifs sur un même thème gouverné par un champ « *in-reply-to* » qui permet de lier les messages entre eux et de reconstituer les échanges à l'aide d'une fonctionnalité présente dans la plupart des clients de messagerie et des *webmails* ;

Gestion électronique de documents (GED) : outil informatisé servant à organiser et gérer des informations et des documents. Les documents ne sont pas forcément nativement numériques dans une GED et peuvent avoir fait l'objet d'une numérisation en entrée ;

JSTOR/Harvard Object Validation Environment (JHOVE) : projet mutualisé entre JSTOR et la Bibliothèque universitaire d'Harvard pour le développement d'une infrastructure scalable de validation de formats. JHOVE est une application écrite en java qui permet de valider les formats spécifiques des objets numériques (Wikipédia) ;

Message électronique : courrier électronique comportant une en-tête, un corps et éventuellement des pièces jointes ;

Messagerie fonctionnelle : boîte attribuée à une entité, un groupe d'individus, qui y ont tous

accès, peuvent lire, répondre aux messages ;

Messagerie interpersonnelle : boîte attribuée à un individu unique qui est responsable de l'utilisation de sa messagerie ;

Messagerie officielle/d'autorité : boîte attribuée à une personne morale ou à un service pour communiquer avec une boîte d'un autre service sans que des personnes physiques y soient individuellement reconnues et identifiées ;

MétaAIP : reprise du concept de l'AIP au sein du modèle conceptuel de la solution VITAM. Le méta-AIP est l'ensemble de fichiers et des métadonnées associées conservées au sein du système d'archivage électronique. Il peut contenir un ou plusieurs versements d'archives et donc un ou plusieurs AIP (Programme VITAM) ;

Métadonnées : ensemble de données caractérisant d'un point de vue technique, structurel et contextuel un document ou une donnée et fournissant l'information indispensable à sa gestion, son accessibilité et son exploitation ;

Metadata Encoding and Transmission Standard (METS) : norme de codage et de transmission des métadonnées en XML, ouvert, non propriétaire, modulaire et extensible, qui permet d'encapsuler plusieurs blocs de métadonnées pour décrire un objet numérique ;

Open archival information system (OAIS) : norme ISO 14721. Norme fondatrice de l'archivage électronique et développé par le centre des études spatiales. Elle définit les grands principes qui régissent un système d'archivage électronique de l'entrée dans le système jusqu'à la génération de versions de diffusion ;

Pièce jointe : document joint à un courriel qui peut être de nature diverse (photo, vidéo, base de données, documents bureautiques, courriel, etc.) ;

Pointeur : référence symbolique à un objet, par exemple une URI ;

Progiciel : logiciel applicatif générique, libre ou exclusif ;

Protocole réseau : ensemble de procédures strictes permettant d'échanger des données entre serveurs sur le réseau informatique (protocole de communication) ;

Protocole IMAP (Internet Message Access Protocol) : il permet une synchronisation de l'état des courriers entre plusieurs clients de messagerie. Les messages restent stockés sur le serveur tout en pouvant être dupliqués sur l'équipement local. Il s'agit du protocole utilisé pour les outils mobiles comme les *smartphones* et les tablettes ;

Protocole POP3 (Post Office Protocol) : les messages sont relevés du serveur et directement stockés sur l'équipement de communication (*device*) de l'utilisateur ;

Protocole SMTP (Simple Message Transfert Protocol) : permet d'envoyer les messages du serveur de transfert (*Mail Transport Agent, MTA*) de l'expéditeur à celui du destinataire. Une fois ce transfert effectué, le serveur MTA du destinataire envoie alors le message au serveur de

messagerie entrant ou MDA (*Mail Delivery Agent*) qui stocke le courrier jusqu'à ce que le destinataire vienne le relever ;

RFC (*Request for comments*) : séries de règles spécifiques développées par l'Internet Engineering Task Force (IETF) pour normaliser les messageries. La plus connue est la RFC 2822 qui spécifie le format des en-têtes et du corps du message, ainsi que les règles d'en-têtes générales comme des champs « To : », « From : » et « Date : » ;

Sémantisation : interprétation du contenu d'un document par l'indexation et la description ;

Serveur de messagerie entrant (*Mail Delivery Agent, MDA*) : serveur temporaire qui stocke le courrier entrant jusqu'à ce que le destinataire vienne le relever ;

Serveur de transfert (*Mail Transport Agent, MTA*) : serveur temporaire de stockage pour transfert *via* internet d'un courriel de la messagerie de l'expéditeur à la messagerie du destinataire ;

Spam ou pourriel : courriel électronique envoyé de manière répétée et non sollicitée sur des messageries électroniques par un expéditeur inconnu du destinataire ;

Standard d'Échange de Données pour l'Archivage (SEDA) : Modélisation des différentes transactions qui peuvent avoir lieu entre des acteurs dans le cadre de l'archivage de documents ou données, s'accompagnant d'une modélisation de la description des données qui seront échangées lors de ces transactions. Le standard propose des schémas XML pour la mise en œuvre de ces transactions fixant la forme des messages échangés ainsi que la forme de la description des données échangées. Il est en cours de normalisation à l'AFNOR (future norme NF Z-44022) ;

Standard MIME (*Multipurpose Internet Mail Extension*) : il étend le format de données des courriels pour leur permettre de supporter des textes en différents codages de caractères et des contenus non textuels. MIME prend en charge le message électronique. Il encode les différentes parties (texte principal, pièces jointes) et place dans l'en-tête les informations pour que le logiciel qui réceptionne le message puisse le décoder. Il définit un ensemble d'attributs additionnels aux en-têtes pour le type de contenu et le codage, ainsi que les règles spécifiques pour encoder les caractères non ASCII dans les en-têtes. Il a remplacé le RFC 2821 qui réglementait auparavant le contenu des messages ;

Submission information package (SIP) : paquet d'information à verser. Ce sigle désigne des archives formatées pour être compatibles avec le système d'archivage électronique et versées dans le système ;

Technologie NoSQL : nouvelle catégorie de base de données qui n'est plus relationnelle mais qui est paramétrable et peut s'adapter à la gestion de volumes importants de données très variables. Elle est utilisée par les grands opérateurs de l'internet comme Google, Amazon ou Facebook et est liée aux technologies dites du *Big Data* ;

TIC : technologies de l'information et de communication ;

Uniform Resource Identifier (URI) : courte chaîne de caractères identifiant une ressource sur un réseau ;

Webmail : interface web rendant possible l'émission, la consultation et la manipulation de courriers électroniques directement sur le web depuis le navigateur internet. Exemple : Gmail, Hotmail, etc.