

Analyse et recommandations sur l'archivage et le versement de documents issus d'applicatifs métier et de GED vers une future plateforme de pérennisation et d'archivage



**Travail de master réalisé par :
Adrian REY RODRIGUEZ**

**Sous la direction de :
Alain DUBOIS, Archiviste cantonal de l'État du Valais**

Genève, le 15 août 2022

**Sciences de l'information
Haute École de Gestion de Genève (HEG-GE)**

Remerciements

Je tiens à remercier ici toutes les personnes qui m'ont aidé, encouragé et soutenu durant la réalisation de ce mémoire :

Mme Cynthia Schneider, archiviste des tpg et mandante de ce travail, pour toute la confiance accordée, sa disponibilité, ses conseils avisés et son optimisme.

M. Alain Dubois, archiviste cantonal de l'État du Valais et directeur de ce travail, pour sa disponibilité, ses conseils avisés et ses encouragements.

Tous les collaborateurs et toutes collaboratrices des tpg qui m'ont accueilli dans leur lieu de travail et qui ont pris part aux nombreux entretiens qui alimentent ce mémoire, pour leur confiance et leur bienveillance.

Je tiens enfin à remercier profondément ma famille et mes amis pour leurs encouragements, pour les relectures et leur présence à mes côtés.

Résumé

Ce travail de master est réalisé pour les Archives des Transports publics genevois (tpg), acteurs majeurs de la mobilité à Genève. Les Archives des tpg s'efforcent d'appliquer le traitement archivistique de leurs fonds depuis 2013, date à laquelle l'entreprise s'est dotée d'une archiviste. Bien que l'archivage papier est un processus maîtrisé, l'archivage électronique n'est pas encore pris en charge et suscite plusieurs questions auxquelles cette étude cherche à répondre. Une quinzaine d'applicatifs métier et plusieurs disques réseau conservent des documents et dossiers à valeur historique, mais la nature et le cycle de vie de ces archives sont peu connus. Une analyse de l'état des lieux est donc effectuée afin de décrire la gestion de ces archives et en vue de leur extraction et versement dans une plateforme de pérennisation et d'archivage.

La réalisation de ce mandat implique des réflexions théoriques pour documenter les connaissances en archivage électronique et des recommandations pratiques pour adapter la mise en œuvre de solutions aux tpg.

Dans un premier temps, une revue de littérature reprenant les divers aspects de l'archivage électronique est proposée. Celle-ci témoigne de la grande variété d'informations disponibles de nos jours autour de ce sujet.

Dans un deuxième temps, des entretiens semi-directifs sont menés auprès des responsables métier et informatiques de chaque application concernée par le projet d'archivage. L'analyse de cet état des lieux, montre l'importante diversité des types d'applications et des données qu'elles contiennent. Elle révèle plusieurs lacunes dans la connaissance de la volumétrie de la masse documentaire et dans la pratique du renseignement des métadonnées. Si dans certains cas, des moyens sont mis en place pour contrôler les durées d'utilité administratives et légales, pour la vaste majorité des applications, ces contrôles ne font pas partie intégrante des paramètres du système. L'étude attire l'attention sur le fait qu'il n'existe pour l'heure pas d'application garantissant la valeur probante des documents, il n'y a pas d'horodatage certifié ni de signature électronique. Cette analyse rend ensuite compte des moyens d'extraction proposés par les responsables des applicatifs métier et de la difficulté de les rendre automatisables.

Enfin, dans un troisième temps, les recommandations formulent des propositions pour la conservation des archives. Les formats existants aux tpg sont listés et des précisions sont apportées pour renseigner les conversions de formats à opérer afin d'assurer la pérennité des données. Les recommandations définissent également les métadonnées minimales et proposent des ébauches de protocoles de versement pour deux applicatifs métier et un disque réseau.

Mots-clés : Transports publics genevois ; tpg ; archivage électronique ; GED ; plateforme de pérennisation et d'archivage ; extraction ; migration ; formats pérennes ; métadonnées ; versement.

Table des matières

Remerciements	i
Résumé	ii
Liste des tableaux	v
Liste des figures.....	vi
1. Introduction.....	1
1.1 Contexte et problématique	2
1.2 But du travail	4
1.3 Objectifs.....	4
2. Méthodologie	6
2.1 Revue de littérature.....	6
2.2 Choix de la méthode d'enquête.....	8
2.2.1 Guide d'entretien semi-directif.....	10
3. État des lieux	12
3.1 Résultats de l'enquête	13
3.1.1 Gestion du cycle de vie	15
3.1.2 Conservation dans les applicatifs et valeur probante.....	17
3.1.3 Moyens d'extraction	18
3.1.4 Critique des résultats de l'état des lieux	19
4. Recommandations	20
4.1 Formats d'archivage long terme et enjeux.....	20
4.1.1 Formats long terme : Guides et outils de référence	22
4.1.2 Formats long terme : Paramètres et critères	23
4.1.3 Formats des documents dans les applicatifs métier et disques réseau ..	25
4.1.4 Recommandations pour les formats	28
4.2 Métadonnées de description et de pérennisation.....	31
4.2.1 Analyse des métadonnées aux tpg	35
4.2.2 Recommandations pour les métadonnées	37
4.3 Création d'un protocole de versement	38
4.3.1 Méthodologie et propositions	40
5. Conclusion	43
Bibliographie	44
Annexe 1 : Diagramme de Gantt	50
Annexe 2 : Grille de lecture	51
Annexe 3 : Liste des archives à valeur historique et leur emplacement	57
Annexe 4 : Guide d'entretien.....	78
Annexe 5 : Tableau synthétique des entretiens	82

Annexe 6 :	Liste des métadonnées pour l'archivage aux tpg	112
Annexe 7 :	Tableau simplifié PAIMAS	116
Annexe 8 :	Protocole de versement – eDocs	123
Annexe 9 :	Protocole de versement – Isilog	127
Annexe 10 :	Protocole de versement – Disque réseau Q	131

Liste des tableaux

Tableau 1 : Mots-clés de recherche	7
Tableau 2 : Types et noms des applications tpg.....	13
Tableau 3 : Niveaux de sensibilité des données tpg.....	16
Tableau 4 : Formats par applicatif métier tpg.....	25
Tableau 5 : Types des formats listés dans les applicatifs métier.....	27
Tableau 6 : Types de formats dans les disques réseau.....	28
Tableau 7 : Recommandations de formats pérennes	29

Liste des figures

Figure 1 : Efficience en tant que valeur dans la stratégie Cap2030	3
Figure 2 : Nombre de sources par thématiques de la grille de lecture	8
Figure 3 : Cycle de vie selon les tpg	12
Figure 4 : Dates d'introduction des applicatifs métier aux tpg	14
Figure 5 : Sensibilité des données dans les applicatifs métier	17
Figure 6 : Trois axes de l'analyse des formats.....	24
Figure 7 : Exemple des métadonnées minimales aux National Archives of Australia...	32
Figure 8 : Modèle de données PREMIS	34
Figure 9 : Protocole de versement au sein d'OAIS	39

1. Introduction

Les Transports publics genevois (tpg) se définissent comme l'acteur majeur de la mobilité à Genève et dans sa région transfrontalière. L'entreprise est devenue une régie publique en 1977, elle se trouve sous la tutelle du Département des Infrastructures et dispose pour la première fois d'une archiviste en 2013 (Thonney 2020). Les tpg sont proactifs dans le développement du réseau de transport genevois, ils sont au cœur de la planification, de la gestion des multiples lignes et de toutes les tâches qui rendent possible les déplacements quotidiens dans la région genevoise.

L'entreprise compte plus de 2180 collaborateurs dont 285 travaillent dans l'Administration (Transports publics genevois 2022a, p. 18), la masse documentaire qui découle principalement des activités constantes au siège des tpg à Bachet-de-Pesay est donc considérable.

Les tpg s'efforcent dès lors d'assurer la mise en place d'une gestion efficace des archives, cela implique entre autres, l'établissement de règles de gestion documentaire et la finalisation d'un calendrier de conservation.

« Le traitement archivistique des fonds a été initié en 2013, avec l'arrivée d'une archiviste. Il a permis de redécouvrir des documents précieux, qui, s'ils avaient bien été conservés durant plus d'un siècle, n'étaient jusqu'alors pas valorisés. L'archiviste a donc débuté ce travail de reprise de l'arriéré en parallèle d'une sensibilisation des collaborateurs et de la mise en place de règles pour la gestion de l'entier du cycle de vie documentaire. »
(Archives des transports publics genevois 2022)

Alors que l'archivage papier est maintenant un processus connu et pratiqué aux tpg et que des démarches et propositions de valorisation ont été mises en œuvre, notamment à travers le Travail de Master de Delphine Thonney en 2020, des questions se posent quant à l'archivage électronique qui est encore considéré comme un processus moins bien connu.

La dématérialisation a déjà pris place aux tpg et continue de se développer depuis plusieurs années. Par dématérialisation, on comprend la numérisation dans certains cas des documents papier ou la création et l'utilisation courante de formats né-numériques. La constante croissance des TIC¹, la poussée vers une gestion des processus plus efficiente et l'adoption de pratiques de travail dématérialisées, rendent omniprésents les contenus au format numérique. L'un des derniers exemples en date est caractérisé par le « virage digital » dans le milieu des ressources humaines, le Rapport annuel de gestion 2020 (Transports publics genevois 2021a, p. 31) indique que :

« Depuis début 2020, les fiches de salaire et les attestations de salaire ne sont plus envoyées en papier au domicile des collaborateurs et collaboratrices, mais mises à disposition sur un portail RH, dans l'espace personnel de chacun et accessible depuis l'application "tpg connect" (sur Smartphones et depuis le site Intranet des tpg) ».

Les Archives des tpg doivent donc élargir leur champ d'action pour traiter également la documentation numérique, ainsi que d'accompagner le changement des pratiques pour assurer une gestion du cycle de vie des documents de qualité.

¹ Technologies de l'information et de la communication.

1.1 Contexte et problématique

Les archives électroniques des tpg ne sont de nos jours, pas encore traitées de manière systématique, il n'y a pas de moyens de collecte, d'évaluation, de classification, de conservation, de communication ou de mise en valeur de ce patrimoine digital. Il n'y a donc pas de gestion du sort final en vue de la destruction ou du versement vers une plateforme de pérennisation et d'archivage.

Les applicatifs métier² sont très nombreux dans l'entreprise et représentent d'importantes quantités de contenus documentaires puisque plusieurs types d'informations découlant de l'activité des tpg y sont créées ou stockées. Plusieurs disques réseau contiennent également des fichiers ou dossiers pour lesquels il n'y a pas encore de traitement archivistique. Le nombre d'applicatifs métier gérant des documents, des photos ou vidéos est d'abord estimé à une trentaine au début de ce travail, les recherches subséquentes réduisent à 16 la liste des applicatifs métier conservant des archives sans compter les disques réseau.

Les lacunes dans le traitement systématique de la masse documentaire numérique empêchent de poursuivre les buts inscrits dans la Politique de gouvernance documentaire des tpg, à savoir, la maîtrise du cycle de vie des documents produits et la garantie de leur protection et conservation de manière durable (Transports publics genevois 2020). La mise en place d'outils d'archivage électronique permettrait une meilleure efficacité de la gestion documentaire et améliorerait la transparence administrative.

Les tpg sont depuis mai 2022 organisés en six domaines d'activité (DA) chapeautés par la Direction générale (DG) qui est formée par deux unités internes. Avant cette date, l'entreprise comptait sept domaines en incluant Infrastructures et bâtiments (I&B)³ (Transports publics genevois 2022c, p. 36).

La composition de la direction est la suivante :

- DG Direction Générale ;
- DG-ME Management d'Entreprise ;
- DG-SJ Service Juridique.

Les domaines d'activité sont les suivants :

- SIT Systèmes d'Information et Télécommunications ;
- E Exploitation ;
- F&G Finances et Gestion ;
- RH Ressources Humaines ;
- T Technique ;
- MVC Marketing, Vente et Communication.

Les DA sont les principaux producteurs de documents numériques aux tpg, ils sont créateurs et utilisateurs de documents à nature très diverse : documents textuels, données, statistiques,

² Terme qui désigne dans le cadre de ce travail tant les logiciels métier que les systèmes GED, outils collaboratifs ou site web et intranet.

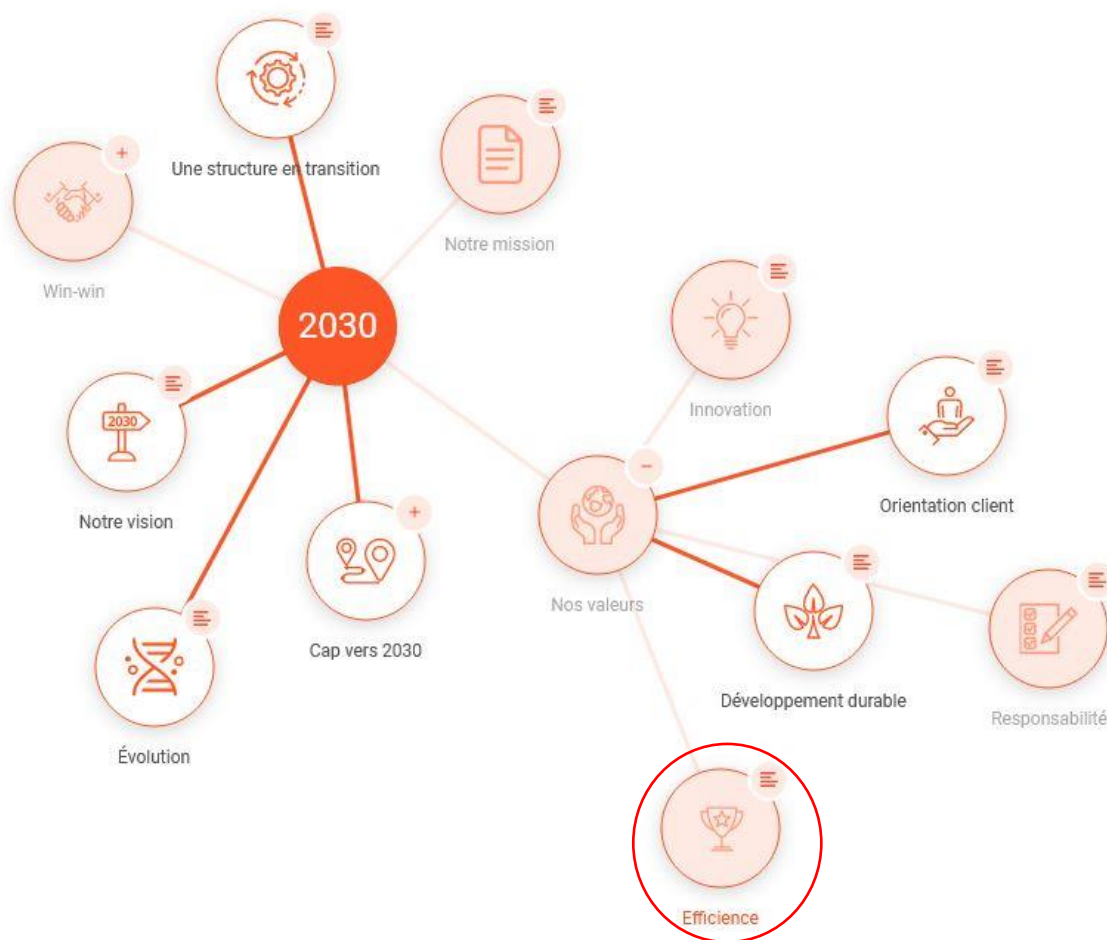
³ Au mois de mai 2022, Infrastructures et Bâtiments est dissout et les équipes de ce domaine sont réparties dans Technique et Marketing, Vente et Communication.

photographies, etc. Plusieurs applicatifs métier dont une GED et des disques réseau sont utilisés dans le cadre de la gestion courante et intermédiaire des informations créées. Ces logiciels et autres éléments informatiques assurent le bon fonctionnement et la continuité des activités des tpg, ils gèrent selon les cas, des documents né-numériques et/ou numérisés.

Cependant, il n'existe pas d'archivage du contenu conservé dans ces applicatifs métier, il n'y a donc pas de gestion du sort final en vue d'un versement aux Archives ou à la destruction. En conséquence, il y a peu de connaissances sur ce qui est produit numériquement par chaque DA et il n'y a pas de perception claire des pratiques de gestion documentaire menées dans les applicatifs métier. Les formats de fichier existants et les pratiques en ce qui concerne le renseignement des métadonnées sont largement méconnus. Il n'y a pas d'aperçu clair de ce que représentent les archives numériques en termes de volumétrie et, dans certains cas, il ne serait pas aisé de rechercher les documents souhaités. Cela augmente les risques de perte d'informations puisque aucune méthode de pérennisation n'est actuellement en vigueur pour ces documents électroniques.

Dans un cadre plus général, les tpg ont élaboré une stratégie de mobilité pour le Grand Genève (nommée Cap2030) afin de répondre aux évolutions de la région et des technologies d'ici 2030. Bien que la stratégie soit axée sur la mobilité future à Genève, l'amélioration de la gestion documentaire exercée par les Archives pourrait bien trouver une place dans les objectifs d'efficacité annoncés par les tpg.

Figure 1 : Efficacité en tant que valeur dans la stratégie Cap2030



(Transports publics genevois 2022b)

En effet, selon la stratégie Cap2030 (Transports publics genevois 2022b) les tpg souhaitent entretenir une culture de la performance et cherchent à optimiser régulièrement les processus pour améliorer leur fonctionnement. Il en va de même pour les Archives qui cherchent à systématiser le versement des documents et s'intéressent ainsi à la gestion des flux documentaires de l'entreprise. Une meilleure connaissance et gestion des informations permet par exemple, de gagner du temps lors des recherches, de contrôler la volumétrie de la masse documentaire et d'améliorer la productivité (Groupe AGS 2021). Ajoutons que selon Coutaz (2016) :

« la croissance irréversible de l'informatique amplifie les phénomènes de masses, en changeant le cadre d'expression et d'archivage. Leur maîtrise ne peut pas se contenter de la simple accumulation au nom des capacités techniques offertes par les nouvelles technologies. »

1.2 But du travail

Outre la problématique de la gestion des documents et du versement, il s'agit de répondre aux questions sur l'archivage à long terme des documents, des formats pérennes et des métadonnées de conservation.

Le but est d'apporter des solutions ou des éléments de base nécessaires pour développer à terme la gestion et la conservation des archives électroniques aux tpg. Le travail cherche à décrire les processus de cycle de vie des documents, en particulier ceux à valeur historique et à proposer des recommandations sur l'entrée des métadonnées minimales et la conservation des divers formats de fichier.

1.3 Objectifs

Pour atteindre le but de ce travail, les objectifs principaux et secondaires sont les suivants :

1. Analyser l'état des lieux.
 - a. Dresser un inventaire des applicatifs métier en usage dans les domaines d'activité ;
 - b. Explorer les fonctions et l'usage de chaque applicatif métier ;
 - c. Sonder les responsables métier et SIT de chaque application (et éventuellement les répondants documentaires des DA⁴) pour détailler l'usage des applications métier et identifier les processus de gestion des séries de documents et dossiers à archiver ;
 - d. Créer un tableau synoptique contenant les résultats de l'analyse de l'état des lieux.
2. Recommander des mesures de conservation.
 - a. Établir des recommandations sur les formats pérennes et les moyens de conservation à long terme ;

⁴ « Le répondant documentaire est choisi pour la bonne vision documentaire qu'il a de son DA. Il sensibilise les collaborateurs de son DA à l'application des règles de bonne gestion documentaire. Les répondants documentaires forment un réseau sur lequel s'appuient le spécialiste de la gestion documentaire (SGD) et l'archiviste. » (Transports publics genevois 2020).

- b. Établir des recommandations sur l'attribution de métadonnées minimales pour les différents types de documents (informations de description et informations de pérennisation) ;
- c. Établir des propositions pour le transfert des archives électroniques vers une future plateforme de pérennisation et d'archivage et pour la création d'un protocole de versement.

2. Méthodologie

La nature de ce travail est à la fois pratique et théorique, les réflexions sont portées sur le périmètre documentaire des tpg et cherchent à apporter des solutions pour des besoins précis en matière d'archivage électronique.

L'archivage électronique implique plusieurs notions conceptuelles, le projet s'intéresse aux flux de gestion, de versement et à l'intégration des modèles de normes et standards pour la description et la conservation. Afin de réduire l'abstrait, la délimitation du sujet et une vision claire des étapes est nécessaire. Comme le rappellent Rietsch et al. (2010, p. 155) :

« Un projet de dématérialisation ou d'archivage électronique peut facilement être assimilé à une expédition... Le risque premier est en effet de méconnaître ses propres obligations et de s'engager dans un projet surdimensionné, sous-dimensionné ou... carrément hors sujet. Malheureusement, même si l'on s'efforce de le rappeler régulièrement, la liste des projets n'ayant pu aboutir est longue, faute d'un véritable dialogue entre les différents acteurs concernés et d'une bonne définition du projet. »

À la suite des réunions de lancement, un cahier des charges est donc rédigé puis validé, celui-ci introduit le mandat, développe les livrables et la planification du temps avec le diagramme de Gantt en annexe 1.

La phase d'initialisation constitue la première étape, elle inclut la création d'une revue de littérature pour cerner les connaissances et pratiques actuelles autour de cette thématique.

S'ensuit l'état des lieux, une étape essentielle pour se lancer dans le vif du sujet et pour collecter et analyser les données. Cette étape inclut le parachèvement de la liste des applicatifs métier dont la fonction et le cycle de vie documentaire sont à renseigner. Il s'agit de répertorier la population cible pertinente pour collecter ces données, en l'occurrence les référent-e-s métier et SIT de chaque application (cf. 2.2 Choix de la méthode d'enquête). Les données issues de l'enquête sont alors regroupées dans un tableau synthétique.

Enfin, les recommandations et propositions puisent dans la revue de littérature pour apporter des solutions adaptées au contexte du terrain tel que spécifié dans l'état des lieux.

2.1 Revue de littérature

La revue de littérature vise à trouver des références pour appuyer les diverses problématiques de l'archivage électronique envisagées par les tpg. Pour cela, une grille de lecture a été créée, elle est disponible en annexe 2.

La grille permet de classer les sources répertoriées dans une ou plusieurs des six thématiques proposées :

- Notions générales (sources qui traitent de l'archivage électronique dans l'ensemble) ;
- Normes et modèles ;
- Versement ;
- Formats ;
- Métadonnées ;

- Présentation de cas (sources qui présentent la mise en œuvre de l'archivage électronique dans des institutions).

Les sources sont identifiées avec leur titre et l'année de parution (les liens URL des sources sont cliquables sous la colonne *Année*), elles sont décrites de manière brève par leur sujet principal ou leur type, une colonne permet de commenter l'information de manière synthétique. La grille regroupe ainsi un total de 44 références dont des articles de revues professionnelles, des sites web institutionnels, des billets de blogs, des normes et modèles, des livres et des outils.

Les recherches ont été menées en français et en anglais, 11 des 44 références sont en anglais. Afin de lancer les recherches, le tableau des mots-clés ci-dessous a été créé :

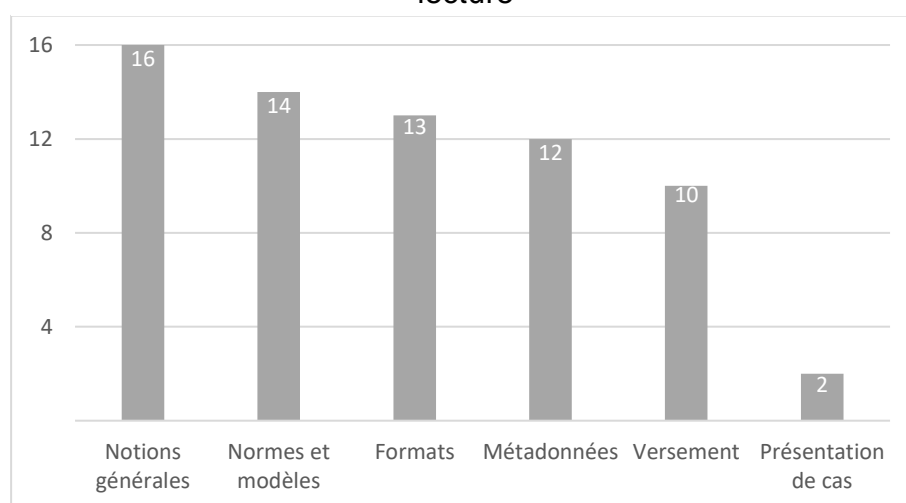
Tableau 1 : Mots-clés de recherche

Français	Anglais
Archivage électronique : vocabulaire général	
Archivage électronique, archives électroniques, archives informatiques, archives numériques	Electronic records, digital records, electronic archiving
Gestion électronique des documents (GED)	Electronic document management system
Gouvernance électronique	Electronic governance
Méthode archivistique, traitement des archives	Archival method
Cycle de vie des documents, cycle de vie documentaire, processus	Records life-cycle, process, processing
Pratiques et moyens	
Normes, modèles, standards	Norms, models, standards
Protocole de versement, convention de versement	Submission Agreement
Outils	tools
Conservation et pérennisation	
Formats pérenns, formats de fichier long terme	Long-term file format, file formats for long-term access, file formats for long-term digital records, file formats for archiving

Métadonnées, métadonnées d'archivage	Metadata, archival metadata
Métadonnées minimales	Minimum metadata, minimum requirements for metadata, metadata for information and records, minimum Mandatory Metadata
Préservation à long terme des données, archivage à long terme	Long-term data preservation, digital preservation, long term archiving

Les requêtes ont été appliquées principalement sur le moteur de recherche Google avec des résultats plutôt satisfaisants et variés tels que le témoigne la grille de lecture en annexe. Il y a cependant dans la liste, peu de sources présentant des cas de projets d'archivage électronique, cela s'explique par le fait que la comparaison avec d'autres institutions n'est pas prioritaire dans le cadre de ce travail qui s'intéresse particulièrement aux spécificités des tpg. Moins d'efforts de recherche ont donc été employés pour documenter ce point. Le visuel ci-dessous reprends le nombre de sources par thématique, notons qu'une source peut couvrir plusieurs thématiques à la fois.

Figure 2 : Nombre de sources par thématiques de la grille de lecture



2.2 Choix de la méthode d'enquête

Le but de l'enquête est de mieux comprendre la fonction des outils listés en annexe 3, au début de ce projet plusieurs outils sont encore méconnus et la forme sous laquelle sont représentés les données à archiver n'est pas toujours évidente. À travers cette démarche, nous souhaitons connaître jusqu'à quel degré l'application est capable de gérer le cycle de vie des documents et quelles sont les pratiques actuelles afin d'identifier les situations où il serait prioritaire d'intervenir. L'enquête cherche également à comprendre les mécanismes et les possibilités ou formes d'extraction pour se préparer au versement dans la plateforme de pérennisation et d'archivage à long terme.

L'enquête aborde donc non seulement des questionnements sur l'utilisation courante, mais aussi des points techniques propres au domaine informatique. Le public cible de l'enquête est constitué de référent-e-s métier et SIT. Les référent-e-s métier ont une vision pratique et une expérience d'utilisation avec les applications concernées par l'archivage. Il s'agit d'un-e ou plusieurs collaborateurs ou collaboratrices désigné-e-s aux tpg comme représentant-e-s d'une ou plusieurs applications et qui connaissent bien leur fonctionnement. Les référent-e-s SIT, ont de leur côté des connaissances techniques, il s'agit d'un informaticien ou d'une informaticienne responsable d'une ou plusieurs applications et qui s'assure de leur maintenance.

La méthode d'enquête s'est tournée vers la conduite d'entretiens semi-directifs avec le public présenté ci-dessus. Les entretiens semi-directifs ont été préférés à d'autres méthodes telles que le sondage en ligne puisque le public est limité et qu'il s'agit d'une étude sur le terrain. La situation des archives électroniques aux tpg est méconnue et peu documentée, la recherche est donc exploratoire et accepte alors des réponses ouvertes. Des observations des pratiques de gestion documentaire au sein des logiciels sont aussi effectuées lors de ces entretiens pour mieux comprendre leur fonctionnement et réduire l'aspect abstrait qui entoure certaines explications. Cette méthode permet une meilleure compréhension du terrain, elle permet d'être réceptif aux possibles remarques et assure une meilleure participation du collaborateur ou de la collaboratrice qui peut recevoir directement des précisions par rapport aux questions posées.

La liste des archives à valeur historique préparée par l'archiviste des tpg en annexe 3, recense les dossiers et documents jugés à valeur archivistique et les rattache à leur applicatif métier ou disque réseau. Cette liste adjoint des remarques et reprend aussi les indications de la DUL (durée d'utilité légale) et du sort final issus du calendrier de conservation.

Pour rechercher les référent-e-s métier et SIT de chaque application listée nous avons fait usage du *Référentiel services et traitements* (Transports publics genevois 2022c). Ce guide fait l'inventaire des services et composants (applicatifs métier), il permet de repérer les principaux responsables applicatifs et entre autres, de situer le niveau de sensibilité des données présentes dans les applications.

Pour organiser les entretiens, une liste de tous les référents métier et SIT a été dressée pour chaque application. Les conseils de l'archiviste ont permis de sélectionner un-e unique interlocuteur ou interlocutrice pour les applications où le référentiel en indique plusieurs.

Afin de fixer les entretiens, des courriels ont ensuite été envoyés au public cible avec les informations nécessaires à la compréhension de la problématique ce qui inclut : un message introductif, la liste des éléments à archiver pour chaque application et un exemplaire du guide d'entretien. La période des entretiens est comprise du 11 avril au 29 juin 2022, cependant cette activité s'est concentrée largement durant le mois de mai. La date du 11 avril correspond à un entretien test qui a permis de valider la faisabilité de la démarche et d'obtenir des premiers retours. La date du 29 juin correspond à une exception faite pour compléter des informations sur un logiciel dont le référent SIT était indisponible durant la période prévue initialement en mai.

Les entretiens ont eu lieu de préférence en présentiel dans les bureaux des tpg et dans certains cas, ils se sont tenus en visioconférence à travers Teams à la demande des

interlocuteurs et interlocutrices. Les entretiens ont en principe été effectués avec un-e seul-e interlocuteur ou interlocutrice à la fois sauf dans deux occasions où deux référent-e-s ont pris volontairement part à l'entretien. Au total, 21 entretiens semi-directifs ont été menés sur place et en ligne.

Les données attendues pour ces entretiens sont de type qualitatif, il s'agit de données narratives. Afin d'illustrer au mieux les réponses données et dans un souci d'éviter les oublis, les entretiens ont été enregistrés et retranscrits dans leur intégralité. Les enregistrements audios ont été supprimés une fois la retranscription achevée. Les réponses des référent-e-s métier et SIT ont été regroupées dans un unique document par application et transmis à l'archiviste des tpg. Un tableau synthétisant les réponses des entretiens a ensuite été créé à partir des retranscriptions (cf. 3.1 Résultats de l'enquête).

2.2.1 Guide d'entretien semi-directif

Le guide d'entretien semi-directif en version générique est disponible en annexe 4. Dans ce guide, les notes en bleu sont des questions de relance, il y a 24 questions au total. La durée d'un entretien est estimée à 40 minutes mais en pratique la durée varie le plus souvent entre 15 et 60 minutes. Cette durée dépend du nombre d'applicatifs métiers interrogés dans un entretien, de la quantité des précisions qui peuvent être apportées par l'interrogé-e et des parties du guide d'entretien retenues pour le ou la référent-e. En effet, le guide est composé de quatre axes dirigés en fonction des situations soit aux référent-e-s métier ou soit aux référent-e-s SIT. Les axes sont les suivants :

1. Profil de l'intervenant et description du logiciel métier ;
2. Cycle de vie des documents/dossiers dans le logiciel ;
3. Conservation dans le logiciel métier ;
4. Extraction des documents/dossiers.

La première partie s'adresse à tous et toutes les interviewé-e-s, elle leur permet de se présenter, de décrire leur activité et l'application. Ces questions cherchent à mieux connaître le profil des représentants des applications afin de garder également une trace de contact utile dans le futur proche lorsque l'archivage électronique prendra forme de manière concrète.

La deuxième partie s'intéresse au cycle de vie des documents et dossiers, dans cet axe nous souhaitons connaître si le logiciel est capable de catégoriser des ressources selon leur statut d'utilisation ou de gérer leur durée de conservation. Nous cherchons également à comprendre si des interconnexions avec d'autres logiciels sont déjà existantes et faciliteraient le transfert des données. Cette partie s'intéresse aussi aux choix de conservation ou d'élimination, aux modalités de recherche au sein de l'application et aux pratiques de renseignement des métadonnées. Les référent-e-s métier sont donc particulièrement visés par les questions de cet axe.

Les troisièmes et quatrièmes parties s'adressent aux référent-e-s SIT. Elles abordent respectivement la conservation et l'extraction.

L'axe de conservation s'intéresse à la garantie de la valeur probante, aux formats des fichiers des archives, leur conservation et volumétrie. En pratique certaines de ces questions ont aussi été posées aux référent-e-s métier, en particulier les questions concernant la valeur probante et les formats.

Enfin, l'axe dédié à l'extraction cherche à prédire les moyens et difficultés pour l'extraction des documents et dossiers listés en annexe 3. Cette dernière partie de l'entretien est particulièrement attentive aux remarques et propositions formulées par les référent-e-s SIT, l'objectif est de comprendre comment extraire de la manière la plus simple possible.

En fin d'entretien une demande de démonstration de l'application est demandée aux référent-e-s métier et/ou SIT afin d'observer les interfaces et d'éclaircir les explications données à propos de l'outil.

Les entretiens ont été adaptés à chaque situation, notamment à travers l'ajout de notes en bleu sous les questions. En effet, les applicatifs métiers sont divers (cf. 3.1 Applicatifs métier des tpg) et certaines questions du guide trop génériques sont donc en décalage avec la réalité de l'outil. Cependant cela n'a pas supposé de difficultés particulières et n'a pas entravé l'aboutissement de l'enquête.

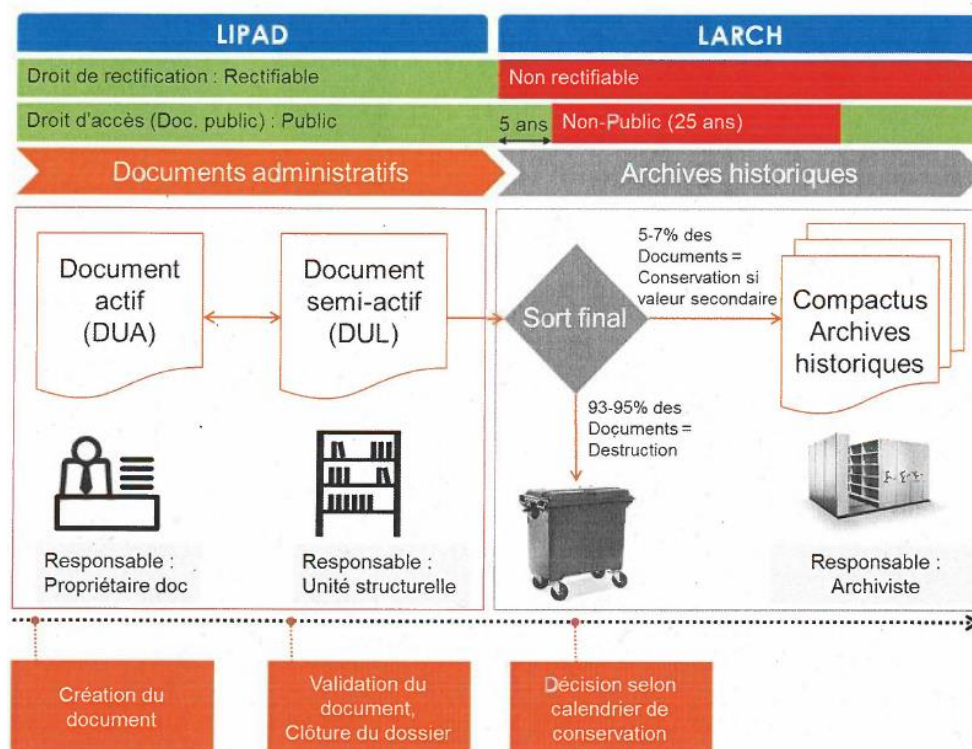
3. État des lieux

Actuellement, les tpg possèdent une politique et un règlement de gouvernance documentaire. La Politique de gouvernance documentaire (Transports publics genevois 2020) s'applique à toute l'entreprise, elle plaide pour un suivi des documents produits et reçus, la gestion efficiente des documents, la conformité aux exigences légales et la protection des documents à valeur probante. Cette politique définit notamment les rôles et les attentes du point de vue documentaire pour les collaborateurs, les directeurs des DA, le responsable d'unité structurelle (d'un service, équipe, etc.), la spécialiste de la gestion documentaire (SGD), l'archiviste et les répondants documentaires. Un point sur la mise en place des outils inclut entre autres, le maintien d'un calendrier de conservation et d'un plan de classement.

Le Règlement de gouvernance documentaire (Schaerer 2021) détaille les règles de bonne gestion dans le but de « soutenir la mise en application de la Politique de gouvernance documentaire. » Ce texte pose le cadre légal et réglementaire auquel les tpg sont soumis en matière de gestion documentaire. En matière de conservation « les tpg s'engagent à gérer la conservation des documents et dossiers tout au long de leur cycle de vie » indépendamment du support, qu'il soit papier ou électronique. Pour les documents électroniques, la responsabilité de l'application du jeu de métadonnée minimal défini par la SGD et l'archiviste remonte à chaque collaborateur. Enfin, notons que :

« Aucun document d'activité ne peut être éliminé définitivement par les collaborateurs-trices sans l'autorisation de l'archiviste. Des exceptions existent pour certains documents (document de travail, papiers de corbeille, brouillons). »

Figure 3 : Cycle de vie selon les tpg



(Schaerer 2021, p. 9)

La mise en œuvre de la politique et la validation et diffusion du règlement sont confiées à la Spécialiste de la gestion documentaire.

La mise en place de ces documents normatifs, est également essentielle pour se préparer à la multiplication des défis futurs de l'archivage dont les composants numériques sont de plus en plus omniprésents.

Les résultats de l'enquête ci-dessous offrent un premier aperçu de l'état des lieux des archives électroniques aux tpg.

3.1 Résultats de l'enquête

L'enquête effectuée à travers plusieurs entretiens semi-directifs a mené à la création d'un tableau synthétique des entretiens, les résultats sont disponibles en annexe 5. Le tableau reprend l'essentiel des réponses pour chaque question en suivant la numérotation originale du guide d'entretien.

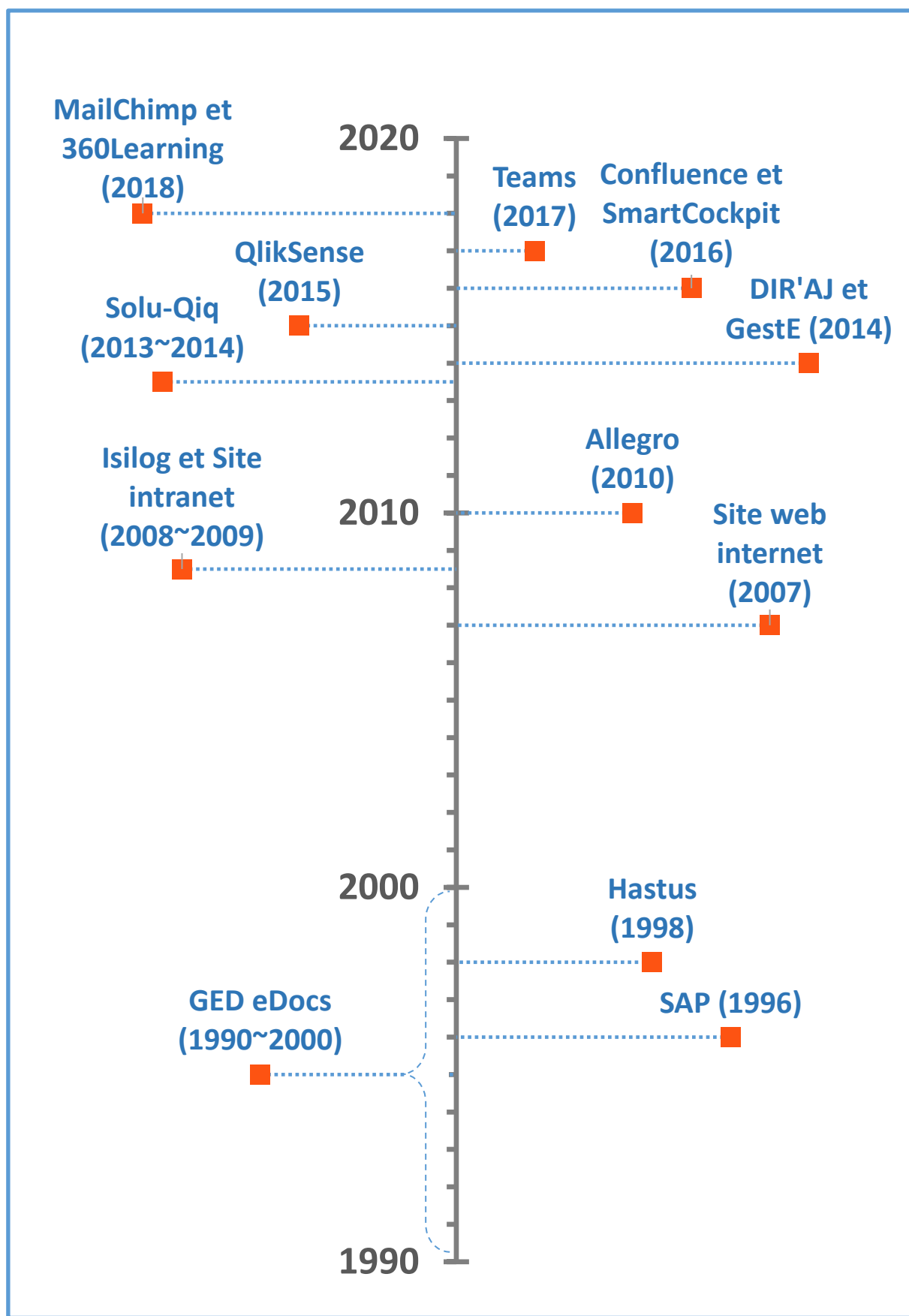
L'enquête détaille les résultats pour un total de 16 applicatifs métier de nature très diverse où se trouvent les archives électroniques listées en annexe 3. Le tableau suivant donne un aperçu de cette diversité.

Tableau 2 : Types et noms des applications tpg

Type d'application	Nom de l'application
Gestion électronique des documents (GED)	eDocs
Enterprise Resource Planning (ERP)	SAP
Learning Management System (LMS)	360 Learning
Logiciels de gestion divers	Allegro (gestion RH) DIR'AJ (gestion juridique) GestE (gestion des événements réseau) Isilog (gestionnaire de suivi d'audits) Hastus (gestion et planification du transport) SmartCockpit (gestionnaire d'indicateurs)
Outils collaboratifs	Confluence (documentation collaborative) Teams
Solution de newsletter	MailChimp
Logiciel analytique	QlikSense
Site web et site intranet des tpg	

Les dates d'introduction de ces applicatifs métier varient des années 1990 à 2018. La GED eDocs serait le plus ancien outil encore en place mais la date précise n'est pas obtenue.

Figure 4 : Dates d'introduction des applicatifs métier aux tpg



On remarque que la décennie allant de 2008 à 2018 est caractérisée par la mise en place de plusieurs nouvelles applications, ce qui marque la tendance vers le tout numérique. Malgré l'ancienneté de certaines applications, leur entretien est assuré par la partie SIT qui effectue régulièrement des sauvegardes et des mises à jour. Toutes les applications présentes dans le visuel sont donc encore en fonction, à l'exception d'Isilog, un gestionnaire de suivi d'audits en phase de décommissionnement. Pour ce logiciel la dernière mise à jour remonte à 2009 et celui-ci ne s'ouvre que sur Internet Explorer, l'ancienne version du navigateur de Microsoft.

3.1.1 Gestion du cycle de vie

Il n'y a pas de pratique uniforme du cycle de vie tant les données, les usages et fonctions des applicatifs métiers sont divers. Plusieurs cas sont possibles en fonction de l'application, un document peut être créé dans un environnement informatique extérieur et importé au sein de l'application ou être créé directement à partir de l'application en question. Il peut aussi s'agir d'un document papier puis numérisé. Dans certains cas, des fichiers joints peuvent alors être rattachés pour compléter le document principal et, s'il y a lieu, des métadonnées sont ajoutées. C'est le cas pour DIR'AJ, l'outil de gestion juridique où les contrats papier sont numérisés, rattachés à des pièces-jointes et où des métadonnées sont entrées manuellement de manière protocolée.

Pour SmartCockpit, gestionnaire d'indicateurs, les données se présentent sous forme de tableaux de bord qui gèrent des indicateurs divers alimentés en temps réel, ce qui permet d'évaluer la performance de l'entreprise dans plusieurs domaines tels que l'économie.

Si l'on s'intéresse à la GED eDocs, les documents doivent s'y retrouver lorsque leur version est finalisée ou en phase d'aboutissement. Les documents sur eDocs peuvent être partagés, modifiés ou consultés. Dès leur validation, la DUA est enclenchée et ils passent plus tard en DUL avant d'être clos.

L'enquête cherche à savoir si les applicatifs métier prennent en charge la gestion des statuts et s'ils indiquent la phase de vie dans laquelle se trouvent les documents (actif, semi-actif ou inactif/version définitive). Les résultats montrent que peu d'applications permettent de classer les documents selon un statut. Pour les quelques applications où cela est possible, le statut d'un document peut être attribué de manière manuelle ou automatique. Sur DIR'AJ, il est par exemple possible de classer manuellement un contrat comme actif ou clos. D'autres assignations similaires sont également possibles avec GestE et Hastus. Du côté de SAP, cette assignation est faite automatiquement mais concerne en l'occurrence le statut de traitement d'une facture. Pour la GED eDocs, cette fonctionnalité devrait être implémentée dans le cadre d'une étude d'analyse et développement.

Il faut toutefois prendre en compte que certaines applications ne peuvent être concernées par l'assignation de statuts, c'est le cas par exemple, pour les sites internet, intranet, le gestionnaire des newsletters MailChimp, l'outil collaboratif Confluence ainsi que d'autres applications où ces champs ne peuvent s'appliquer étant donné la nature des données ou des fonctions du logiciel.

Ces résultats se répètent pour l'assignation des durées de conservation de type DUL. Des systèmes de rappel sous forme d'agenda existent pour DIR'AJ et GestE, mais l'implémentation d'un système gérant concrètement la DUL serait souhaitée. Dans le cas de SAP, la montée de version « S/4 » devrait permettre d'assigner des durées de conservation.

3.1.1.1 Interconnexions, pratiques de gestion, recherche et sensibilité des données

En ce qui concerne les interconnexions qui pourraient faciliter le transfert des données entre les logiciels, les informations fournies n'encouragent pas à mener un approfondissement de cette question. C'est particulièrement le cas avec SAP qui ne serait pas évident d'usage pour les exportations de documents.

En règle générale, lorsqu'un document atteint sa version finale aux tpg, il est conservé indéfiniment, il n'y a pas de traitement particulier. Dans certaines applications, les anciennes versions ou documents finalisés sont retirés de l'affichage mais toujours conservés soit dans l'application même ou ailleurs dans un serveur. On dénombre tout de même quelques exceptions, en effet, sur Solu-Qiq les anciennes versions d'organigrammes sont écrasées par les nouvelles. Dans Hastus, bien que la règle soit de tout conserver et que le logiciel existe depuis 1998, les données ne remontent que jusqu'à 2009. Dans le site internet qui existe depuis 2007, les back-ups ne remontent que jusqu'à 2012. Le visuel du site internet est par ailleurs sauvegardé à la fréquence d'une fois par année depuis 2012 par la Bibliothèque nationale suisse à travers son portail *e-Helvetica Acess*, ce portail conserve les sites suisses d'intérêt historique (Bibliothèque nationale suisse 2022).

Les pratiques de renseignement des métadonnées sont détaillées au point 4.2.1 (Analyse des métadonnées aux tpg). Il n'y a dans la plupart des cas, pas ou très peu d'ajouts de métadonnées aux documents, mis à part certaines métadonnées renseignées automatiquement comme l'horodatage simple. Il n'y a pas de propositions pour l'ajout de champs de métadonnées à renseigner et certain-e-s référent-e-s n'ont jamais eu l'occasion d'aborder cette question auparavant. Néanmoins, les résultats de l'enquête montrent que la recherche des documents suppose peu de problèmes, les outils de recherche proposés par les applications en usage seraient simples et performants. On remarque pourtant que pour GestE, l'outil de gestion des événements réseaux, la réussite d'une recherche est bel et bien dépendante de la qualité du renseignement des métadonnées.

Enfin, au niveau de la sensibilité des données, on dénombre quatre niveaux définis par le tableau suivant :

Tableau 3 : Niveaux de sensibilité des données tpg

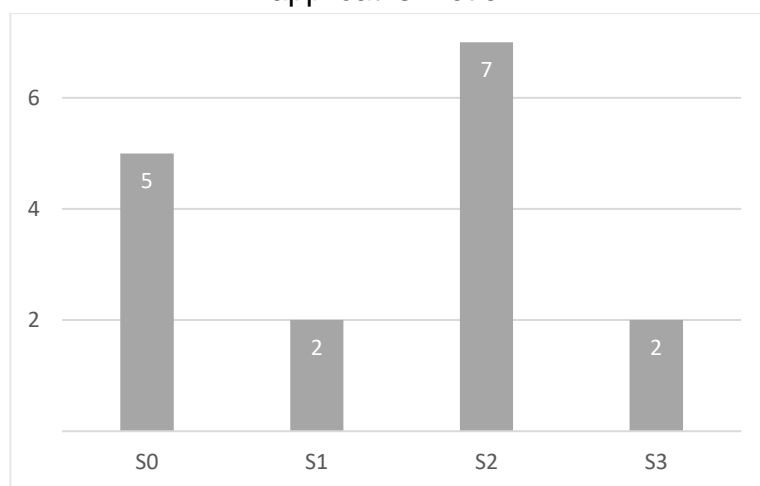
S0	Données personnelles sensibles ⁸ ou données confidentielles	Données médicales, liées aux aides sociales, liées aux poursuites judiciaires Données du conseil d'administration
S1	Données personnelles non sensibles ⁷	Donnée clientèle voyageur
S2	Données internes	informations relatives au fonctionnement des services, informations non classifiées explicitement
S3	Données publiques	information voyageur, communication aux médias

(Transports publics genevois 2022c, p. 3)

Chaque applicatif métier est classé dans un des niveaux S0 pour les données les plus sensibles à S3 pour les données publiques, ces informations reprises du Référentiel services et traitements seraient à jour. Les données présentes dans les documents et dossiers à

archiver ont des niveaux de sensibilité très hétérogènes, comme le montre le visuel ci-dessous, plusieurs de ces archives peuvent comporter des informations de nature personnelle et sensible et doivent faire objet d'une attention particulière afin d'éviter les fuites en cas de d'export et de migration. La plupart des informations sont toutefois à caractère interne, c'est le cas pour au moins 7 applicatifs métier dont la GED eDocs. L'ancien référent SIT d'eDocs fait d'ailleurs remarquer qu'aucune réflexion n'a été menée pour définir le sens sémantique de ces éléments et que la gestion du droit d'accès est au bon-vouloir de l'auteur du document⁵.

Figure 5 : Sensibilité des données dans les applicatifs métier



3.1.2 Conservation dans les applicatifs et valeur probante

Les archives électroniques incluent des documents à valeur probante tels que des contrats, hors l'étude montre que les applicatifs métier ne peuvent garantir la valeur probante de ces documents à un niveau élevé. L'horodatage est courant mais il n'est pas certifié et la signature électronique est inexistante, les contrats au format numérique possèdent toujours leur version originale en papier qui fait foi. Si les paramètres actuels répondent du moins aux exigences internes des tpg, ils ne peuvent répondre aux exigences juridiques du point de vue externe⁶.

Au niveau de la conservation, le détail de l'analyse des formats est disponible au point 4.1.3 (Formats des documents dans les applicatifs métier et disques réseau). Les problèmes liés à l'ouverture des fichiers sont d'habitude inexistantes et les quelques cas de fichiers corrompus ou déformés, sont résolus grâce aux nombreuses sauvegardes et restaurations effectuées par le service informatique.

L'évaluation de la volumétrie des documents et dossiers à archiver est nécessaire pour quantifier les informations à traiter par les archivistes et estimer les coûts de stockage à venir dans la plateforme de pérennisation et d'archivage. Malgré l'inclusion des référents SIT dans les entretiens, les réponses sont peu satisfaisantes et ne permettent pas d'évaluer précisément la volumétrie des données à conserver. En effet, les quelques informations récoltées et disponibles en annexe 5 (à la question 3.4) ne font qu'énumérer le nombre

⁵ Entretien avec l'interviewé 22, Responsable Infrastructures et Opérations, Genève, 11 avril 2022.

⁶ Ibid.

d'éléments totaux dans un logiciel ou le stockage total de celui-ci. Pour obtenir des renseignements plus précis, les interrogé-e-s proposent de se rediriger vers d'autres collaborateurs plus en lien avec les serveurs ou vers les ingénieurs en infrastructures. De manière globale, on pourrait néanmoins s'attendre à quelques voir à plusieurs centaines de gigabytes à archiver dans l'immédiat. Pour les disques réseau, la volumétrie est également difficile à estimer puisque l'accès à plusieurs dossiers est restreint. Pour le disque réseau Q on peut tout de même estimer environ 11 GB de données à archiver (cf. annexe 10).

Malgré le fait que la recherche d'informations dans les applicatifs métier serait plutôt aisée comme indiqué précédemment au point 3.1.1.1, certain-e-s référent-e-s ont exprimé de la peine à retrouver les documents et dossiers à partir des informations listées en annexe 3. Cette difficulté s'explique par la présence de documents issus d'autres domaines d'activité dont le ou la référent-e ne fait pas partie. Dans le cas de la GED eDocs, c'est l'absence d'indication de l'identifiant numérique qui empêche les recherches.

3.1.3 Moyens d'extraction

Les modalités d'extraction des documents ou dossiers et de leurs métadonnées sont des informations clés pour envisager la manière de connecter la plateforme de pérennisation et d'archivage. En ce sens, les Archives souhaitent des recommandations et des retours sur l'existence d'une API (Application Programming Interface) afin de lancer des requêtes d'extraction.

Ici encore, les réponses sont diverses, plusieurs moyens d'extractions sont possibles en fonction des particularités de chaque application mais des doutes persistent. L'utilisation d'une API serait envisageable pour eDocs et Confluence. Cependant, pour une majorité d'applications les référent-e-s opteraient pour une extraction manuelle. C'est un choix qui se justifie facilement pour des applications avec peu d'éléments à extraire comme MailChimp qui ne compte qu'une douzaine d'infolettres. Pour d'autres structures comme SmartCockpit et Hastus où plusieurs éléments devraient être extraits, la tâche manuelle serait conséquente.

Les référent-e-s SIT de SmartCockpit, GestE et Allegro proposent d'ailleurs de reléguer la tâche d'extraction aux fournisseurs de ces applications. Dans le cas d'Allegro, gestionnaire des APOP, une extraction à partir de requêtes SQL est aussi possible, mais l'affichage des informations ne serait plus sous forme de documents textuels mais sous forme de tables dans une base de données. Des choix sont donc à faire, tout en veillant à tenir compte de la nature, de l'intégrité des informations et des contraintes techniques de l'extraction.

On évoque aussi des outils tels que des aspirateurs de site web⁷ pour l'extraction des visuels du site internet, intranet et de Solu-Qiq qui contient des organigrammes cliquables.

Enfin, quelques difficultés sont à prévoir lors de l'extraction puisque certains formats de documents pourraient subir des altérations lors des manipulations de conservation. Cela concernerait les documents tels que les fichiers Word, PowerPoint ou d'autres types de fichiers qui ne sont pas au format PDF, celui-ci est considéré comme plus stable.

⁷ « Un aspirateur de site Web est un type de logiciel qui télécharge toutes les données contenues par un site Web pour les sauvegarder sur un support de mémoire local de l'utilisateur, ce qui permet de consulter les pages correspondantes sans être connecté à Internet. » (Aspirateur de site web 2019).

3.1.4 Critique des résultats de l'état des lieux

Les entretiens semi-directifs ont permis de mettre à jour des informations issues du *Référentiel services et traitements* concernant les responsables des applicatifs métier et d'éclaircir le rôle de ces outils de gestion. Cette étude a aussi permis d'élucider des incertitudes quant à l'existence de certaines applications initialement retenues dans la liste des archives historiques puis retirées, il s'agit de :

- BDD Carto qui est en réalité un terme générique pour *Base de données cartographique* et dont les plans d'exécutions sont enregistrés dans le disque réseau H ;
- CMDB (Configuration Management Database), un logiciel permettant d'inventorier les composants informatiques et dont l'implémentation est en cours de discussion dans l'environnement SIT ;
- DAO (Design assisté par ordinateur) constitué de plans réalisés sur AutoCAD mais présents dans le disque réseau Q ;
- Isiline qui n'est pas un logiciel mais le nom du fournisseur d'Isilog.

Deux autres applications sont listées en annexe 3 mais ne font pas partie des enquêtes : tpgConnect et OneNote. Les entretiens prévus pour tpgConnect ont été annulés et remplacés par des échanges de courriels directement avec l'archiviste. L'outil OneNote a été ajouté à la liste des archives plus tardivement et n'a donc pas été inclus dans l'étude.

Cependant, les résultats obtenus au fil des entretiens sont constitués d'informations peu précises, ce qui rend difficile l'analyse et la comparaison des résultats tant le niveau de précision des informations fournies varie selon l'application étudiée.

4. Recommandations

Le transfert de données des outils métier vers la plateforme de pérennisation et d'archivage implique plusieurs préparatifs en vue de la conservation et du bon déroulement des opérations :

- La migration des fichiers dans des formats stables et avec des perspectives de conservation optimales ;
- La définition des métadonnées qui accompagnent les documents ;
- La préparation des protocoles de versement afin de résumer les processus de l'extraction au versement.

Notons qu'au cours de cette année, les Archives des tpg ont choisi la solution d'archivage électronique docuteam cosmos⁸ de l'entreprise suisse docuteam SA.

Les recommandations se basent à la fois sur des éléments théoriques et sur les résultats de l'enquête menée dans ce travail. Les recommandations s'intéressent notamment au Modèle de référence pour un Système ouvert d'archivage d'information (OAIS) qui aborde les divers aspects de l'archive électronique (Consultative Committee for Space Data System 2012, p. 1.2) :

« Le Modèle de référence aborde toute une série de fonctions liées à la pérennisation d'informations archivées comprenant le versement des données, le stockage, la gestion de données, l'accès et la diffusion. Il aborde également la migration d'informations numériques vers des supports et des formats nouveaux, les modèles de données utilisés pour représenter l'information, le rôle du logiciel dans la pérennisation de l'information, ainsi que l'échange d'informations numériques entre Archives. Il identifie les interfaces internes et externes des fonctions d'archivage ainsi qu'un certain nombre de services de haut niveau relatifs à ces interfaces. »

4.1 Formats d'archivage long terme et enjeux

L'une des attentes pour ce travail, est de spécifier les formats de conservation pour les divers types de contenus qui seront versés dans la future plateforme de pérennisation et d'archivage. Il s'agit d'abord de lister les formats existants actuellement dans les applicatifs métier, puis d'analyser leur adéquation en vue d'un archivage long terme. Lors du versement des archives dans la plateforme de pérennisation et d'archivage, les formats qualifiés comme non-adaptés à la conservation à long terme devront être convertis vers des formats qui garantissent mieux la conservation et leur lecture dans le temps.

La question du choix des formats pour l'archivage est importante, car elle conditionne la manière dont les archives seront gérées et préservées. C'est ainsi que la page web du Queensland Gouvernement (Australie), consacrée aux recommandations sur les formats numériques introduit l'importance du sujet : « The file format you choose will affect how records are preserved and managed. The choice of format becomes more critical the longer a record has to be kept. » (Queensland Government 2021).

La notion de « long terme » demeure vague, selon le Modèle de référence pour un Système ouvert d'archivage d'information OAIS (Consultative Committee for Space Data System 2012), le long terme est une :

⁸ Lien vers la page web docuteam cosmos : <https://www.docuteam.ch/fr/docuteam-cosmos/>

« période de temps suffisamment longue pour qu'il soit nécessaire de prendre en compte les changements technologiques qui ont un impact sur l'information détenue par un OAIS, et notamment la gestion des nouveaux supports et formats de données ainsi que l'évolution de la Communauté d'utilisateurs cible. Cette période n'est pas limitée dans le temps. »

Dans l'idéal un format long terme devrait donc être en mesure de conserver l'information souhaitée de manière illimitée dans le temps, mais cela semble irréaliste car tout comme pour les supports de stockage, les formats sont amenés à évoluer. Deux articles publiés dans *Les petites cases*, un blog consacré aux technologies du web, laissent réfléchir quant à la pertinence de qualifier certains formats comme étant pérennes :

« Aucun format n'est pérenne, car il est impossible d'imaginer toutes les évolutions possibles. Il est vrai, en revanche, que certains formats de par leurs caractéristiques sont plus simples à gérer sur le long terme que d'autres. » (Got [Pseudonyme], 2010)

« Il n'existe pas de support de stockage pérenne. Certains présentent plus d'assurance que d'autres comme c'est le cas des bandes magnétiques LTO, mais sans garantie absolue » ... « Un format n'est pérenne que si l'organisation qui conserve est capable d'en maîtriser aussi l'exploitation sur le long terme. » (Got [Pseudonyme], 2018)

Garantir la pérennité d'un format reviendrait à s'assurer régulièrement que les paramètres qui le définissent comme format pérenne sont constamment à l'ordre du jour.

Les Archives de l'État de Berne (2022), décrivent aussi l'archivage numérique comme un processus continu :

« L'archivage numérique est un processus continu jamais achevé qui nécessite sans cesse de nouvelles mesures. Les documents sont conservés dans un petit nombre de formats de fichiers pérennes adaptés à l'archivage, qui s'accordent aux conditions environnementales à mesure qu'elles se modifient. Ainsi, l'information est conservée et peut être utilisée même à de nouvelles conditions. »

Duranti (2021), professeure d'archivistique et directrice du projet InterPARES (International Research on Permanent Authentic Records in Electronic Systems) résume de manière très simple la question de la pérennité des documents numériques : « It is not possible to preserve a digital records. We can only preserve our ability to reproduce or recreate it. »

Contrairement aux supports analogiques comme le papier où l'information (le texte) et le support (le papier) ne forment qu'un seul et unique élément à conserver, le numérique est plus complexe : l'information est contenue dans un format qui peut changer, le format contenant l'information est stocké et consultable à travers un support lecture dont le fonctionnement est lui aussi sujet au passage du temps et donc à l'obsolescence etc. L'Australasian Digital Recordkeeping Initiative (ADRI 2020), explique ci-dessous la complexité de la conservation du numérique en exemplifiant les diverses raisons qui pourraient faire qu'une image au format JPEG ne soit plus lisible au bout d'une longue période :

« A particular challenge of digital information is that it is necessary to use software to access the information held in a digital object. It is necessary to use an image viewer to display a JPEG image, for example. Accessing digital information means having software to open the digital object for the entire lifespan of the information. For example, it may not be possible to find software that displays JPEG images in 100 years if the JPEG format has fallen out of use. If software to access a format is not available, information stored in this format will become inaccessible. »

Dunant Gonzenbach (2013) met en évidence la nécessité d'un entretien régulier des données numériques pour assurer leur existence sur le long terme et s'appuie sur l'exemple ci-dessous :

« La conservation de l'information sous forme numérique est beaucoup plus complexe que la conservation sous forme papier ou microfilm. Les documents électroniques ont une grande vulnérabilité au temps et ont pour caractéristique les formats et supports les moins durables de l'histoire. Un cd par exemple a une durée de vie d'une dizaine d'années. On peut affirmer aujourd'hui qu'un objet numérique créé sans précautions particulières ne pourra plus être lu d'ici une quinzaine d'années. » ... « La question des formats est critique, car elle constitue un point particulièrement difficile à résoudre dans la perspective de la conservation à moyen et à long terme de l'information. Les évolutions techniques sont si rapides qu'il est difficile de présager la solution qu'il faudra adopter. C'est pourquoi il est raisonnable de conserver les données dans des formats répondant à des critères définis. »

Au vu des textes cités dans ce chapitre, il est manifeste que la conservation des formats passe par une révision régulière de leur condition de lecture et environnement technologique, les paramètres qui caractérisent la sûreté d'un format sur le long terme doivent être définis et mis à jour de manière continue.

4.1.1 Formats long terme : Guides et outils de référence

De nombreuses institutions ont établi des politiques et guides de référence afin de normaliser la pratique et d'aider les archivistes et professionnels du domaine documentaire à faire des choix de formats cohérents et durables. En Suisse, nous pouvons tout d'abord citer le Centre de coordination pour l'archivage à long terme de documents électroniques (CECO 2022a), dont la mission est de « soutenir les Archives membres dans toutes les questions que leur pose l'archivage de documents numériques. » Le CECO émet à travers de son site web, des recommandations de formats de données aptes à l'archivage et décrit plus de 52 formats sur la base d'analyses et d'évaluations. Nous pouvons aussi citer le guide de normes et standards émit par les Archives fédérales suisses (AFS 2020), celles-ci spécifient les formats acceptés par les Archives lors des propositions de documents par les services. Au niveau genevois, les Archives d'État publient elles aussi depuis 2011 des recommandations adressées à l'administration et une liste des formats obligatoires pour l'archivage électronique.

En Europe, nous pouvons mentionner la Bibliothèque nationale de France qui a mis en ligne en 2021 une version initiale pour appel à commentaires de sa politique en matière de formats de données pour la préservation à long terme. La politique aborde non-seulement les recommandations pour le choix de formats pérennes, mais aussi les métadonnées qui peuvent accompagner les divers types de documents.

Une autre initiative intéressante se situe au Royaume-Uni où les Archives nationales (The National Archives) mettent à disposition en ligne un registre technique nommé PRONOM, il s'agit d'une base de données avec laquelle il est possible d'obtenir de l'information technique sur les formats existants, entre autres, la version, la description, le développeur et la date de sortie. La base de données comporte plus de 1'300 formats dans son registre, elle est mise à jour régulièrement (The National Archives 2022). Il est aussi possible d'obtenir un identifiant unique PUID (PRONOM unique identifier) pour se référer de manière inéquivoque à une version précise d'un format. PRONOM met aussi à disposition des ressources tels que des outils comme DROID (Digital Record Object Identification), cet outil permet d'identifier automatiquement les formats dans des lots de fichiers.

À l'international, nous pouvons mentionner dans l'espace nord-américain, les lignes directrices de la Bibliothèque et Archives Canada (BAC 2015) sur les formats de fichiers et la page web *Sustainability of Digital Formats* du site de la Library of Congress (2022a) aux États-Unis. Enfin, en Océanie, le Queensland Government (2021) en Australie met à disposition une page web synthétisant les principaux points à tenir en compte pour l'évaluation de la durabilité des formats. Le site pointe aussi vers un document plus complet : le *Sustainable Digital File Formats* mis au point par un sous-comité de l'Australasian Digital Recordkeeping Initiative (ADRI 2020).

4.1.2 Formats long terme : Paramètres et critères

Les formats long terme ou pérennes sont qualifiés ainsi parce que des paramètres ou critères reposant sur divers facteurs en rapport à leur forme de création, utilisation, documentation et accessibilité, garantissent une meilleure gestion et un minimum d'assurance face aux risques d'obsolescence.

Les guides et politiques mentionnés plus haut (cf. 4.1.1 Formats long terme : Guides et outils de référence), conseillent tout d'abord de se limiter à un choix restreint de formats pour l'archivage. Cela dans le but de diminuer la complexité de leur gestion (par exemple : la gestion des mises à jour), d'uniformiser l'information pour la rendre plus lisible et de diminuer les coûts d'entretien. Les Archives d'État de Genève (2011a), considèrent que :

« Ce choix de limiter les formats d'archivage est effectué dans toutes les institutions pratiquant de l'archivage électronique. En effet, un petit nombre de formats soigneusement sélectionnés et contrôlables garantit de manière plus sûre leur lisibilité sur le long terme qu'une grande quantité de formats dont l'entretien sera plus complexe et onéreux. »

Dans la même ligne, les Archives fédérales suisses (2020) disent privilégier la qualité à la quantité :

« Un nombre clair et restreint de formats soigneusement choisis est nettement plus à même de garantir la compréhension ultérieure qu'un grand nombre de formats difficilement contrôlable, ces formats demandant en outre beaucoup d'entretien et dépendant souvent des versions actuelles des applications informatiques. »

Selon le Queensland Government (2021), il faut se prévenir des formats qui ne sont plus pris en charge par leur développeur, qui utilisent une compression avec pertes (c'est à dire que la compression du fichier mène à une perte d'informations en parallèle de la diminution de sa taille) ou qui sont restreints par la propriété intellectuelle (par exemple : des formats sous licence et qui ne partagent pas librement les spécifications techniques).

Au contraire, les formats long terme sont :

- De nature libre et sont documentés selon des standards mis en place par des organismes de normalisation et ne dépendent pas de matériel informatique spécifique ou d'un logiciel pour être lus. La documentation attachée au format est importante pour l'utilisateur afin que celui-ci puisse éventuellement remettre en marche un logiciel capable de lire le format dans le futur (Australasian Digital Recordkeeping Initiative 2020, p. 29) ;
- Courants : l'utilisation active et courante d'un format par la population ou une communauté (notion d'ubiquité) est un signe de durabilité : « plus la communauté d'utilisateurs est conséquente, plus elle est susceptible d'avoir développé ou fait développer des outils adaptés. » (Bibliothèque nationale

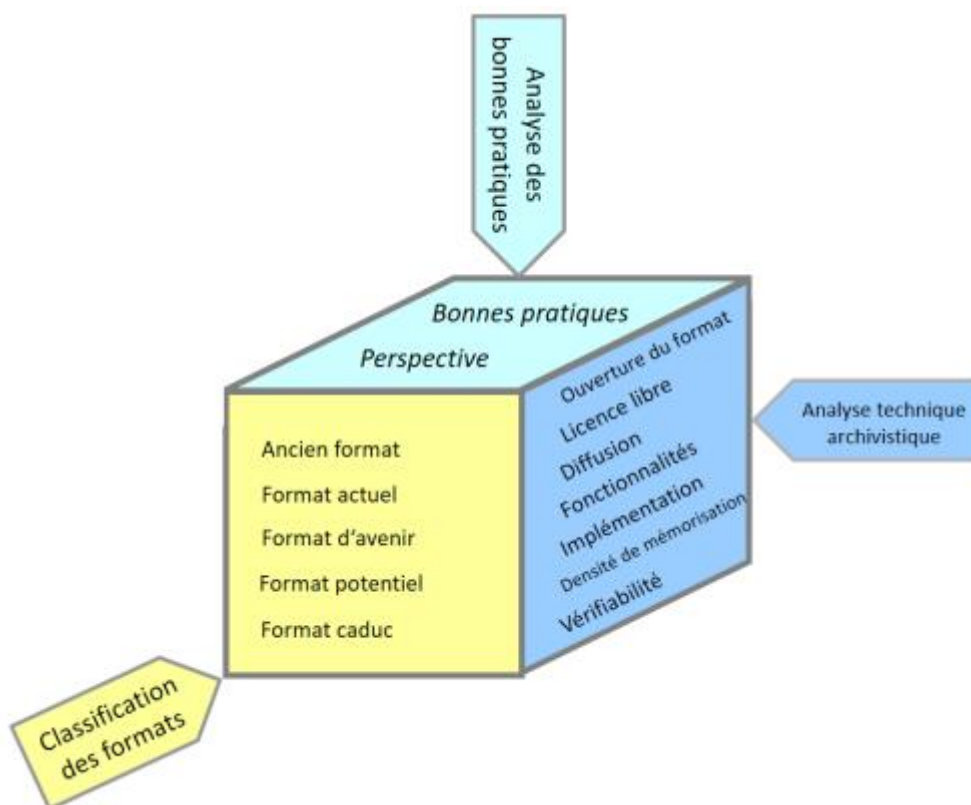
de France 2021, p. 12). Il s'agit de formats très répandus et qui obligent les logiciels à les lire sous peine de ne pas être intéressants ou pratiques à utiliser, cela inciterait donc les développeurs à entretenir les dits formats (Australasian Digital Recordkeeping Initiative 2020, p. 28) ;

- Stables : peu de besoins de mises à jour et de changements vers des nouvelles versions ;
- Non-compressés ou utilisent une compression sans perte : les informations contenues dans un fichier compressé ne subissent pas de perte et le fichier peut être remis dans son état original ;
- Rétenteurs de métadonnées : le fichier contient lui-même des métadonnées, ce qui facilite la compréhension de son contexte et la recherche d'informations (Australasian Digital Recordkeeping Initiative 2020, p. 30).

Cette liste ci-dessus n'est pas exhaustive et reprend les principaux points permettant de se faire une première idée de la durabilité d'un format. L'Australasian Digital Recordkeeping Initiative (2020, p. 28) propose huit critères pour ce type d'évaluation et la Bibliothèque nationale de France (2021, p. 11) en propose quatorze : douze sont qualifiés de critères objectifs et deux de subjectifs. Les critères subjectifs « reposent sur une évaluation directement liée aux besoins et aux moyens de l'utilisateur réalisant l'évaluation. »

Enfin, le CECO se base sur trois axes de critères pour analyser l'adéquation des formats au long terme comme le montre la figure ci-dessous :

Figure 6 : Trois axes de l'analyse des formats



(Centre de coordination pour l'archivage à long terme de documents électroniques 2021)

C'est à partir de l'analyse technique archivistique et des bonnes pratiques que le CECO propose une classification des formats en fonction de leur aptitude au long terme.

4.1.3 Formats des documents dans les applicatifs métier et disques réseau

Les formats issus des applicatifs métier aux tpg sont en nombre restreint. Malgré certaines difficultés des intervenants pour lister de manière exhaustive les formats pour chaque applicatif métier, les documents PDF et Word (DOCX) sont respectivement les plus répandus dans la liste des applicatifs métier analysés. Les formats renseignés pour chaque applicatif métier sont à lire comme suit dans le tableau plus bas :

- Format suivi de l'exposant « p » comme probable : Ce format est probablement présent dans l'applicatif métier, il n'y a pas de données précises à ce sujet.
- Format suivi de l'exposant « j » comme joint : Ce format est présent dans l'applicatif métier sous forme de pièce-jointe.
- Format suivi de l'exposant « e » comme export : Ce format n'apparaît qu'à l'export d'une page ou d'un élément contenant de l'information à partir de l'applicatif métier.
- Format suivi de l'exposant « e+p » : Ce format peut probablement être envisagé à partir de l'export.

À noter qu'à plusieurs reprises les indications quant aux formats gérés par les applicatifs métier sont peu précises, certain-e-s référent-e-s interrogés ont sommairement mentionné que le ou les logiciels en question contiennent des documents et formats de la suite Microsoft Office (Microsoft 365). Dans le tableau ci-dessous les extensions de la suite Microsoft Office sont suivies de (x), puisque l'extension exacte n'a pas pu être vérifiée. En effet, si l'on prend l'exemple d'un fichier Word, l'extension peut être DOC si la version est antérieure à 2007 ou DOCX pour les versions actuelles (Crider 2017).

Tableau 4 : Formats par applicatif métier tpg

Applicatif métier	Formats
GED eDocs	PDF, DOC(X), PPT(X), XLS(X) et tout autre format ⁹
SAP	PDF, XLS(X), CSV, HTML ^{e+p}
Hastus	PDF, JPEG ^p
Site web internet	PDF, JPEG ^p , PNG ^p , MP4 ^p
Isilog	PDF ⁱ , DOC(X) ^j , PPT(X) ^j , XLS(X) ^j ou autres formats de la suite Microsoft Office ⁱ , SQL ^{e+p}
Site intranet	PDF
Allegro	RTF ^e , SQL ^{e+p}

⁹ eDocs ne tient pas en compte le format du fichier, cette plateforme peut tout accepter.

Solu-Qiq	PDF, DOC(X), XLS(X) ou autres formats de la suite Microsoft Office, HTML ^{e+p}
DiraJ	PDF, DOC(X) ^j , MSG ^j
GestE	PDF, DOC(X), XLS(X), JPEG, PNG ^p , MSG
QlikSense	PDF ^e , XLS(X) ^{e+p} , JPEG ^{e+p}
Confluence	PDF ⁱ , DOC(X) ⁱ , VSD(X) ^j ou autres formats de la suite Microsoft Office ^j , « Wiki ¹⁰ »
SmartCockpit	PDF ^j ou formats de la suite Microsoft Office ^j
Teams ¹¹	PDF ^p , DOC(X), PPT(X), XLS(X) ou autres formats de la suite Microsoft Office
MailChimp	PDF ^e , MSG (copie MSG séparément sur le client de messagerie Outlook)
360 Learning	PDF ^j , PPT(X) ^j ou autres formats de la suite Microsoft Office, JPEG ^j , MP4 ^j

On note ici qu'à l'exception d'Allegro, logiciel qui gère les APOP (évaluations des collaborateurs), tous les applicatifs contiennent du PDF et/ou peuvent exporter les informations dans un format PDF. Les documents DOC(X) seraient de la même manière, présents dans 9 des 16 applicatifs métier.

Tous les formats sont connus et largement répandus même parmi les moins courants de la liste tels que le RTF, MSG, HTML, SQL ou MP4. Un doute subsiste pour le format dit « Wiki » sur Confluence, logiciel de documentation collaboratif fonctionnant sous le modèle de Wikipédia, car le format n'est pas reconnu en tant que tel par la base de données PRONOM. Selon deux articles présents sur Wikipédia :

« Un wiki est une application web qui permet la création, la modification et l'illustration collaboratives de pages à l'intérieur d'un site web. Il utilise un langage de balisage et son contenu est modifiable au moyen d'un navigateur web. » (Wiki 2022)

« Le wikitexte ou wikicode est un langage de balisage léger qui définit la mise en forme de saisies de contenu utilisateur. Il est le plus souvent utilisé pour écrire les pages d'un wiki... Certains logiciels wiki mêlent une syntaxe similaire à celle de certaines balises du langage HTML et une syntaxe propre. » (Wikitexte 2022)

Nous pouvons donc inclure le format dit « Wiki » dans la catégorie des textes balisés.

¹⁰ Les pages de Confluence seraient au format texte « Wiki » selon le référent SIT interrogé.

¹¹ À noter qu'il n'y aura pas d'extraction à partir de Teams puisque celui-ci n'est pas cadré et qu'il s'agit d'un outil de stockage à court terme.

Notons que pour 360 Learning, il serait possible de créer les formations dans le format SCORM (Sharable Content Object Reference Model), un standard pour les cours en ligne et interopérable pour les plateformes *eLearning* (SCORM 2022). Les formations créées sous le format SCORM, peuvent être exportées à nouveau en ce même format¹². En ce moment il n'est possible d'extraire uniquement les pièces-jointes des formations, tandis qu'avec SCORM il serait possible d'extraire la formation en entier. SCORM n'est cependant pas en usage aux tpg.

Le tableau ci-dessous, catégorise les différents formats listés dans les applicatifs métier en types d'informations. Pour classer les divers formats en catégories, le guide de l'Australasian Digital Recordkeeping Initiative (2020, p. 7) a été consulté, il présente en effet, un exemple de classification des formats par types d'informations. La base de données PRONOM renseigne aussi l'information de classification dans les métadonnées du format recherché, elle a également été consultée pour parfaire les indications du tableau.

Tableau 5 : Types des formats listés dans les applicatifs métier

Types	Formats
Documents de traitement de texte, présentations ou publications	PDF, DOC(X), PPT(X), RTF
Set de données, tableaux, feuilles de calcul	XLS(X), CSV
Images vectorielles (représentations et diagrammes)	VSD(X)
Texte balisé, (site web)	HTML, « Wiki »
Texte structuré	SQL
Messagerie	MSG
Images	JPEG, PNG
Vidéos	MP4

Notons que dû à la difficulté, lors des entretiens, de vérifier exhaustivement les formats composant les documents listés pour l'archivage historique, il se peut que certains formats listés ne fassent pas partie des documents à archiver.

Concernant les disques réseau, l'exploration des dossiers et séries de documents listés pour l'archivage nous permet de constater des formats plus spécifiques. Ces éléments sont répartis sur six disques réseau avec des niveaux d'accès différent. Deux des six disques réseaux, les disques P et Q (annexe 3) sont accessibles mais peuvent contenir des dossiers dont l'accès est restreint, raison pour laquelle la liste des formats dans les disques réseau peut s'avérer

¹² Entretien avec l'interviewé 2, Ingénieur d'application IT, Genève, 29 juin 2022.

incomplète. Parmi les éléments qui ont pu être observés les dates extrêmes vont de 2001 à 2022.

Les formats des disques réseau sont dans l'ensemble semblables aux formats listés pour les applicatifs métier, les documents de traitement de texte, les sets de données et tableaux sont identiques. Il y a une importante présence de photographies au format JPEG (plusieurs centaines) et en moindre mesure au format TIFF. La majorité des documents sont également au format PDF et DOC(X)

Certains formats sont spécifiques aux activités des tpg, il y a ainsi plusieurs DWG, un format natif du logiciel AutoCAD utilisé dans le Design assisté par ordinateur (DAO) pour les plans, ces fichiers peuvent contenir des images matricielles BMP. Il existe des vidéos au format MP4 mais les dossiers explorés et accessibles n'en contiennent pas.

Tableau 6 : Types de formats dans les disques réseau

Types	Formats
Documents de traitement de texte, présentations ou publications	PDF, DOC(X), PPT(X), RTF, TXT
Set de données, tableaux, feuilles de calcul	XLS(X)
Images vectorielles (représentations et diagrammes)	DWG, DXF, IDW
Texte balisé (site web)	HTML
Messagerie	MSG
Images	JPEG, TIFF
Vidéos	MP4 ^p

4.1.4 Recommandations pour les formats

Au premier abord, les formats listés dans les applicatifs métier sont largement répandus et connus du public. Plusieurs formats font partie de la suite Microsoft Office et il y aurait peu ou pas de fichiers lourds tels que des vidéos et images à extraire des applicatifs.

Lors des entretiens, il a été mentionné la possible utilisation de requêtes SQL pour extraire des informations sous forme de tables de données pour les logiciels Allegro et Isilog. SQL serait dans ce cas, probablement le moyen d'extraction mais pas le format de conservation.

Pour les disques réseaux, les images sont beaucoup plus présentes ainsi que les images vectorielles telles que le DWG.

Les recommandations de conservation pour les formats, sont établies à l'aide des guides et outils de référence cités au point 4.1.1 de ce travail. Les recommandations ci-dessous, synthétisent les démarches pour chaque type de format et renseignent l'identifiant PUID du format à conserver.

Tableau 7 : Recommandations de formats pérennes

Types	Formats d'origine	Format d'archivage (PUID)
Documents textes basés sur pages	PDF, DOC(X), PPT(X), RTF	PDF/A-1 (fmt/95) ou PDF/A-2 (fmt/476)
Documents textes non structurés	TXT	Pas de conversion.
Set de données, tableaux, feuilles de calcul	XLS	XLSX (fmt/214)
	XLSX	Pas de conversion.
	CSV	SIARD (fmt/161)
Images vectorielles	VSD	VSDX (fmt/924)
	DWG, DXF, IDW	Pas de conversion.
Images	JPEG	JPEG2000 (x-fmt/392) ou TIFF (fmt/353)
	TIFF	Pas de conversion.
	PNG	Pas de conversion.
Vidéos	MP4	FFV1 (fmt/569) ou pas de conversion.
Messagerie	MSG	Pas de conversion.
Texte balisé	HTML	Pas de conversion.
	« Wiki »	PDF/A-1 (fmt/95) ou PDF/A-2 (fmt/476)

4.1.4.1 Recommandations pour les documents texte

Pour les documents texte basés sur des pages, comme les documents PDF, DOC(X), PPT(X) et RTF la conversion en PDF/A est largement présentée comme la solution pour une

conservation long terme. Il y a deux versions de PDF/A qui se prêtent à l'archivage long terme : le PDF/A-1 est spécifié dans la norme ISO-19005-1 et le PDF/A-2 dans ISO-19005-2. Le CECO (2021a) « recommande la version PDF/A-2 au niveau de conformité U pour la création de nouveaux documents PDF/A afin de permettre la transparence et de garantir la facilité de recherche. » Le PDF/A-2u évite en effet une erreur apparue au niveau de la conformité B qui peut rendre impossible la recherche plein texte ou son extraction (CECO 2022b).

Le PDF/A est auto-documenté, il intègre « tous les éléments nécessaires à la restitution du document, et notamment les polices de caractère et les informations colorimétriques dont il a besoin. » (Archives d'État de Genève 2011a, p. 5).

Les documents texte non structurés, comme le format TXT n'ont pas besoin d'être convertis, TXT est considéré comme un format très stable.

4.1.4.2 Recommandations pour les feuilles de calcul et set de données

Les tpg contiennent dans les disques réseau et probablement dans certains applicatifs métier, des documents XLS en parallèle des XLSX. XLS est un format propriétaire de Microsoft Excel précédant XLSX, le format est périmé et les spécifications de sa structure sont détenues par Microsoft (CECO 2021b). Bien que XLSX ne soit pas spécifié comme un format d'archivage, il peut être envisageable de convertir les XLS vers la version plus récente XLSX. En effet, selon le CECO (2021c) : « Etant donné qu'aucun format d'archivage ne se profile vraiment jusqu'ici dans le domaine des tableurs, de plus en plus d'institutions acceptent cependant XLSX au moins comme format d'archivage provisoire. »

Le CSV est un format de données structurées, il est amplement utilisé même dans l'archivage, le format n'est pas normalisé et selon le CECO (2021d), il est nécessaire : « une documentation précise sur le masquage des séparateurs et des retours automatiques à la ligne, l'utilisation de guillemets et le problème de la variation du nombre de champs par ligne. » La conversion des CSV dans un codage XML résoudrait les problèmes mentionnés ci-dessus. Il serait envisageable de convertir les CSV en format SIARD (une norme eCH), les Archives fédérales suisses (2020, p. 13) recommandent de convertir les bases de données relationnelles en SIARD et proposent l'outil *csv2siard*.

4.1.4.3 Recommandations pour les images vectorielles

Les informations à propos de VSD(X) sont limitées parmi les sources retenues, si des anciennes versions VSD existent il serait possible de les mettre à jour en VSDX qui est développé et documenté par Microsoft (Library of Congress 2021).

Les documents DWG et DXF sont des formats propriétaires de la société Autodesk, DXF est toutefois ouvert. Les deux formats ne sont pas aptes à l'archivage long terme, ils sont dépendants du programme AutoCAD fourni par Autodesk, selon le CECO (2021e et 2021f) il n'y a pas d'alternatives.

IDW fait également partie d'AutoDesk mais il n'y a pas d'informations à propos de son archivage, IDW peut être exporté en DWG (FileInfo 2021).

4.1.4.4 Recommandations pour les images et vidéos

Les images JPEG sont nombreuses dans les disques réseaux, certaines sont au format TIFF et il pourrait y avoir des images PNG dans les applicatifs métier.

JPEG est déconseillé à l'archivage car sa compression entraîne des pertes qui pourraient s'aggraver lors de migrations, les images de type photographique devraient être converties dans un format plus adapté au long terme. Le CECO (2021g) avertit que ce format :

« Convient mal aux graphiques linéaires... Si les fichiers graphiques livrés aux archives se trouvent déjà en format JPEG, la migration en formats sans pertes n'a de sens que si ceux-ci présentent de meilleures perspectives à long terme... Lorsque des fichiers graphiques importants sont fournis dans ce format, il est préférable de renoncer à une migration plutôt que de convertir dans un format d'archivage qui ne convient pas. »

À noter que la conversion de fichiers JPEG en TIFF ne solutionnerait pas les pertes de données ayant déjà eu lieu dans le format JPEG.

Le format TIFF est accepté par les Archives d'État de Genève (2011a, p. 7) et les Archives fédérales suisses (2020, p. 5), ce format est largement accepté et sa norme est ouverte.

PNG est aussi un format ouvert et un standard ISO, le CECO (2021h) le considère apte à l'archivage mais conseille JPEG2000 pour les images de type photographique.

Dans le cas où des vidéos au format MP4 devraient être archivées, les Archives fédérales suisses (2020, p. 5) recommandent la conversion au format FFV1 qui est conçu pour l'archivage. Il convient de s'assurer de la volumétrie des vidéos, si la qualité de l'image n'est pas prioritaire, le format MP4 (pratique pour la diffusion en ligne) peut être conservé tel quel.

4.1.4.5 Recommandations pour les autres formats

Peu d'informations couvrent les e-mails au format MSG (messagerie Outlook) parmi les sources consultées. MSG est documenté mais est un format propriétaire de Microsoft (Library of Congress 2020). Ce format est accepté par la Bibliothèque et Archives Canada (2015).

Concernant HTML, le CECO (2021i) spécifie que :

« Du fait de la séparation de la structure des contenus et de la présentation formelle, le format répondrait en fait aux besoins de base de l'archivage numérique, puisqu'il rendrait l'archivage indépendant du développement du logiciel de visualisation. Par contre, le développement très rapide de la technologie des navigateurs a modifié entretemps le format et il s'avère que le contenu de beaucoup de pages internet n'est compréhensible que si la présentation graphique est correcte. »

La même source indique que HTML5 présenterait un potentiel dans le développement d'un format adapté à l'archivage, mais qu'il ne peut pas encore être recommandé.

Les pages « Wiki » du logiciel Confluence sont exportables en PDF directement depuis l'appliquatif métier, leur conversion en PDF/A est donc à envisager.

SQL n'est pas repris dans le tableau des recommandations de formats pérennes (cf. 4.1.4). Selon le CECO (2021j), SQL n'est « pas d'un format de données mais d'une méthode permettant de créer une base de données relationnelle à partir d'un fichier de *scriptage* (fichier texte) et réciproquement. » Pour l'archivage de bases de données il est recommandé d'utiliser SIARD.

4.2 Métadonnées de description et de pérennisation

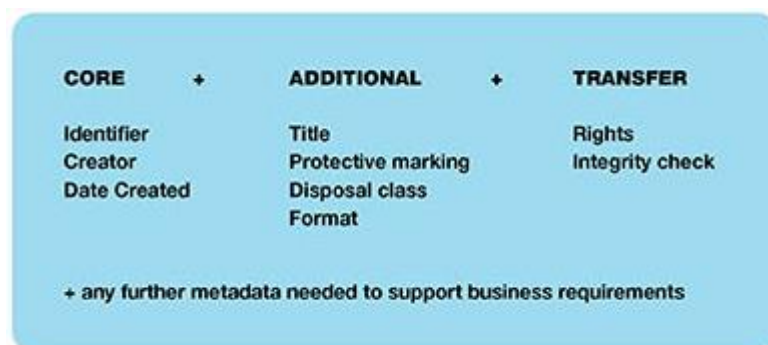
Les métadonnées sont des données sur des données, par exemple, des informations associées à un fichier image, texte etc. Les métadonnées servent à trouver les informations

nécessaires et à comprendre plus facilement les documents, elles permettent d'authentifier le contenu, la provenance et à comprendre l'usage du document en question (Archives New Zealand 2022a). Ces informations rendent donc plus accessibles les documents car elles définissent leur contexte de création et d'utilisation, elles génèrent ainsi un meilleur environnement pour le partage et l'interopérabilité des systèmes au sein d'une organisation.

Les métadonnées font l'objet de normalisation dans nombre d'institutions qui les définissent en catégories et précisent les informations minimales requises pour une documentation correcte des données.

Les Archives New Zealand (2022b) et les National Archives of Australia (2022) ont mis en ligne leurs exigences en matière de métadonnées pour leurs administrations respectives. Ces exigences définissent les métadonnées à renseigner durant la phase de création du document et celles à attribuer lors de l'évaluation pour le versement aux Archives. Les National Archives of Australia (2022) perçoivent trois catégories d'informations : le *Core*, c'est-à-dire les métadonnées principales pour tout document ; les métadonnées additionnelles pour les documents à plus haute valeur et conservation sur le long terme, et enfin, les propriétés de transfert qui renseignent les droits d'accès.

Figure 7 : Exemple des métadonnées minimales aux National Archives of Australia



(National Archives of Australia 2022)

Les métadonnées peuvent aussi être séparées en deux types : les informations de description et les informations de préservation. Les normes existantes offrent des schémas de métadonnées génériques qui permettent de mieux préciser le contexte des documents et d'utiliser un vocabulaire contrôlé pour décrire les informations. C'est le cas de Dublin Core Metadata Initiative (DCMI), organisation qui œuvre pour les bonnes pratiques en matière de métadonnées. Celle-ci propose une liste de 15 champs de description génériques qui forment un standard (DCMI Usage Board 2012 ; Bibliothèque nationale de France 2008, pp. 3-16) :

- Titre ;
- Créateur ;
- Sujet ;
- Description ;
- Éditeur (entité responsable de la mise à disposition ou diffusion) ;
- Contributeur ;
- Date ;

- Type ;
- Format (identification du moyen logiciel ou matériel pour accéder à la ressource et/ou dimensions de la ressource, par exemple : le nombre de pages) ;
- Identifiant de la ressource ;
- Source ;
- Langue ;
- Relation (Références aux ressources apparentées) ;
- Couverture (domaine d'application de la ressource) ;
- Gestion des droits (restrictions d'utilisation etc.).

Si l'on s'intéresse plus précisément à la description en milieu archivistique, il convient de se familiariser à EAD (Encoded Archival Description), un modèle XML basé sur la norme ISAD(G) et maintenu par la Library of Congress (2022b). Ce modèle possède un équivalent pour la norme ISAAR(CPF) sur les notices d'autorité : EAC-CPF (Encoded Archival Context).

EAD permet de « traiter les instruments de recherche existants, de restituer l'organisation hiérarchisée des instruments de recherche et de conserver le principe d'héritage des informations entre les niveaux » (Bibliothèque nationale de France 2022b). La comparaison des éléments EAD (Society of American Archivists 2004) avec les éléments descriptifs de la norme ISAD(G) (Conseil international des archives 2000) permet de constater la similitude des deux textes. Nous pouvons donc retrouver comme dans ISAD(G), les éléments des zones d'identification, de contexte, de structure, etc.

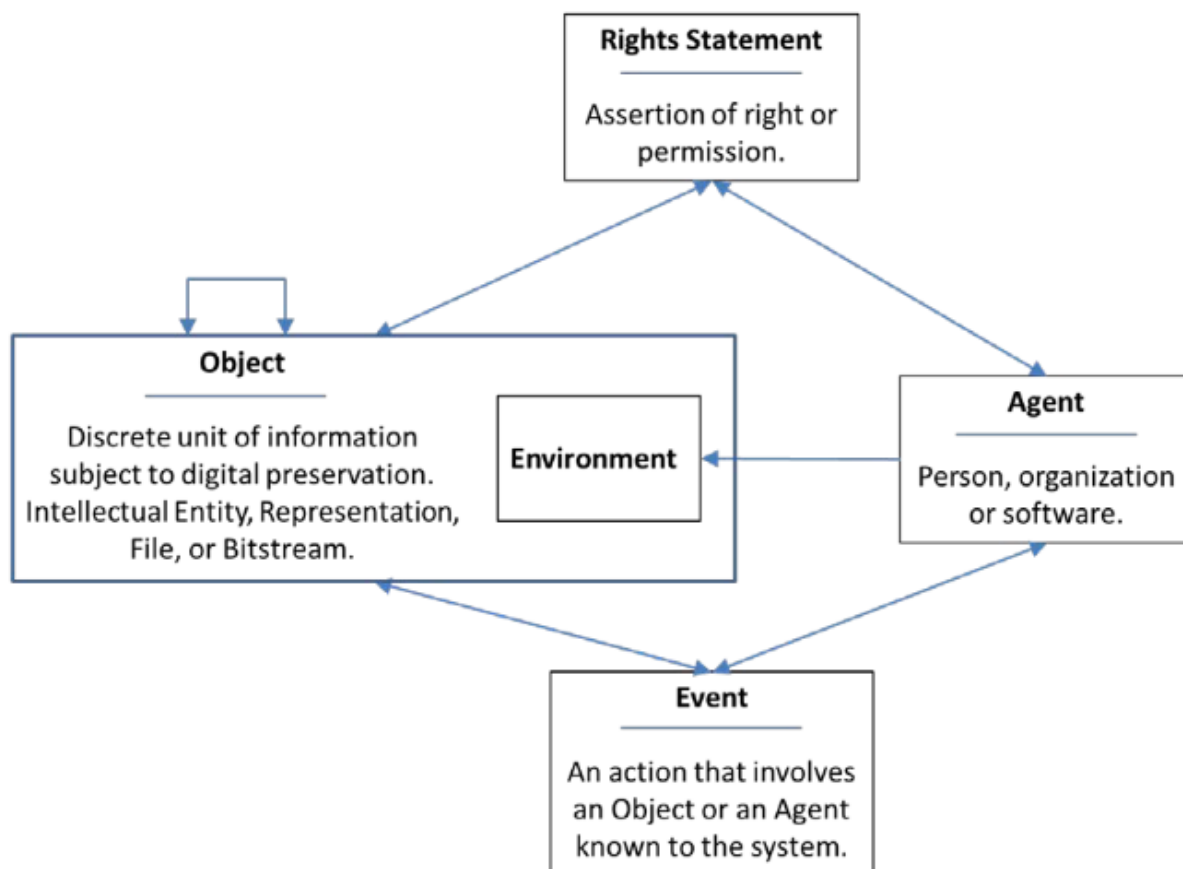
En ce qui concerne les informations de pérennisation, PREMIS (PREservation Metadata Implementation Strategies) est le standard pour la préservation des documents et objets numériques. PREMIS se présente sous la forme d'un dictionnaire de données, il est aussi maintenu par la Library of Congress. Ce standard « définit les métadonnées nécessaires en général pour remplir les fonctions de préservation sur tout type de collection » (Caplan 2009, p. 5).

Selon Caplan (2009, p. 3), les métadonnées de préservation sont utiles pour « garantir l'usage à long terme d'une ressource numérique. » Caplan cite quelques exemples pour situer l'intérêt des métadonnées de préservation :

- Ces métadonnées permettent de tracer l'historique des modifications en enregistrant l'empreinte d'un fichier ;
- Elles définissent le support physique dans lesquels les fichiers sont stockés, ce qui permet d'être attentif aux évolutions et obsolescences des supports afin d'opérer les migrations nécessaires ;
- Les informations sur les formats et leur environnement logiciel permettent de garder la trace des formats originaux et des programmes qui les lisent ce qui permet aussi de mener les opérations de préservation à ce niveau ;
- Elles documentent la provenance numérique du fichier, ses détenteurs et les interventions à travers le temps permettent d'authentifier le document même après des modifications.

Le modèle de données PREMIS met en relation quatre entités qui doivent être décrites : les objets, les événements, les agents et les droits :

Figure 8 : Modèle de données PREMIS



(PREMIS Editorial Committee 2015)

Dans la figure ci-dessus, l'objet est la ressource à préserver, il s'agit de l'unité d'information, le fichier ou document. Il peut aussi s'agir d'une représentation, un dossier constitué d'un ensemble de fichiers. Un environnement, c'est-à-dire le support informatique de lecture ou un logiciel, peuvent aussi être des objets à préserver.

Les événements représentent les opérations de préservation sur un objet, ces événements sont menés par des agents qui sont des personnes, organisations ou logiciels. Ceux-ci peuvent attribuer des droits ou autorisations sur un objet ou sur les agents.

Les droits impliquent les objets et les agents, ils se matérialisent sous forme de « propriété intellectuelle (copyright), texte de loi (statute), convention avec des ayants droit (license) » (Bibliothèque nationale de France 2022a).

Comme l'indique Caplan (2009, pp. 10-12), chacune de ces entités comporte des informations à renseigner et chaque entité a un identifiant unique. L'entité Objet comporte les données relatives à son intégrité, son format, sa taille, les informations sur sa création, son environnement (son support de stockage et ses relations avec d'autres objets) et les restrictions.

Pour renseigner les événements, il faut d'abord définir l'action de l'évènement, s'il s'agit d'une création, d'un versement ou d'une migration. Il faut situer l'évènement dans le temps et décrire les résultats.

L'agent doit être identifié par son nom et par sa nature : personne, organisation ou logiciel.

Enfin, les droits sont définis par des informations détaillées sur leur statut, les actions qu'ils autorisent ou leurs limitations, il faut renseigner leur durée de validité, les objets et les agents impliqués par la déclaration de droits.

Les Archives d'État de Genève (2011b) listent sous forme de tableau les champs de métadonnées obligatoires et facultatifs à implémenter dans une application de type GED. Ce tableau est apparu dans le cadre du projet Gal@atae, un projet d'archivage électronique à très long terme entre 2011 et 2013 (Dunant Gonzenbach 2013). Les AEG spécifient si ces métadonnées doivent être reprises ou non lors du versement aux Archives et indiquent les types de champs pour chaque cas : Numérique, alphanumérique, liste déroulante et format de date. Ce document a été repris et est en cours d'adaptation dans le cadre de la Gouvernance documentaire des tpg, son champ d'application s'étend aux domaines d'activités (DA) ainsi qu'au Conseil d'administration (CA) où les documents sont créés, reçus, utilisés et conservés (Schaerer, Schneider et Thonney 2022).

4.2.1 Analyse des métadonnées aux tpg

Le tableau du document cité ci-dessus (cf. 4.2), a été repris pour son adaptation aux tpg, il est disponible en annexe 6. À cette date, le document n'est pas encore validé, cependant il est attendu que les référents des services informatiques tpg intègrent ces métadonnées dans les applicatifs métier dont la GED. Le jeu de métadonnées des tpg s'applique à tous les systèmes qui contiennent des documents et qui répondent aux critères minimaux tels que définis dans le Règlement de gouvernance documentaire tpg (Schaerer 2021, p. 5).

En règle générale, toutes les métadonnées renseignées par les producteurs avant l'évaluation seraient conservées en cas d'archivage. Le tableau en annexe 6 détaille vingt champs de métadonnées dont treize sont obligatoires (O) et sept sont facultatifs (F). Parmi ces champs, onze métadonnées peuvent être reprises lors du versement aux Archives si tous les champs facultatifs correspondants sont renseignés.

Dans un scénario de renseignement minimal où seules les métadonnées obligatoires seraient reprises à l'archivage, nous retrouverions les huit éléments suivants :

- Nom du producteur ;
- Intitulé du dossier ;
- Rubrique du plan de classement ;
- Type de document ;
- Sous-type de document ;
- Intitulé du document ;
- Date de création du document ;
- Numéro unique.

Certains des champs qui ne sont pas retenus pour l'archivage, n'ont un intérêt que durant la phase de vie active et semi-active du document, ils concernent par exemple, le numéro de règle du calendrier de conservation et les informations sur la DUAL (durées d'utilité administratives et légales).

Actuellement, les métadonnées sont encore sous-exploitées dans la plupart des applicatifs métier. Il existe un écart conséquent entre des applicatifs ne fonctionnant que si tous les champs obligatoires ou exigés par les besoins du métier sont remplis et les applicatifs pour lesquels il n'existe pas de telles règles et où le renseignement détaillé n'est pas une condition pour le bon fonctionnement du service.

En effet, pour des applications comme DIR'AJ, GestE, Isilog ou SAP, il est exigé de renseigner plusieurs champs d'informations. On remarque qu'il s'agit d'applications d'une grande importance pour les tpg : DIR'AJ est l'outil de suivi juridique, il contient des dossiers de contrats et des documents de nature probante ; GestE gère les tâches liées aux sinistres et accidents, il permet de monitorer les interventions ; Isilog a pour fonction le suivi des audits d'entreprise et permet de créer des fiches d'amélioration ; enfin, SAP est la solution ERP qui permet de gérer une multitude de processus métier.

Dans chacun de ces quatre applicatifs métier, la plupart des métadonnées sont entrées manuellement. Dans le cas de DIR'AJ, la formation d'un dossier avec contrat passe par le remplissage protocolé d'un formulaire de données concernant le contrat, il y a aussi un menu déroulant pour indiquer l'état du cycle de vie par rapport à la DUAL. Il n'y a pas de métadonnées générées automatiquement en dehors des métadonnées automatiques classiques comme l'horodatage d'un document ou le nom du créateur. Il n'a pas été possible d'obtenir de détails des métadonnées entrées manuellement.

Lors de la réception d'un constat sur GestE, les parties prenantes sont, entre autres, renseignées avec les noms, prénoms, adresse et assurance. Une liste indique les champs à remplir, si les éléments ne sont pas correctement enregistrés cela peut compliquer la recherche des dossiers¹³. L'horodatage et le nom du créateur du dossier sont automatiquement remplis.

La création d'une fiche sur Isilog entraîne l'entrée de plusieurs informations : la source de la demande (audit interne ou externe), le responsable de l'action et, entre autres, le responsable SQSE (Santé, Qualité, Sécurité, Environnement). Les métadonnées obligatoires sont signalées par Isilog par un astérisque, le logiciel ne fonctionne que si tous les champs sont renseignés.

En finance, les champs obligatoires recensés pour SAP sont : la date de pièce, la date comptable, la devise, le type de pièce (une facture fournisseur, un client, un débit, un crédit), la clé de comptabilisation, la nature comptable, le code TVA, le montant et l'information analytique (information d'un ordre ou d'un centre pour la comptabilité analytique).

D'autres applications exploitent les métadonnées dans une moindre mesure et à des niveaux de rigueur différents.

La GED eDocs peut récupérer automatiquement les informations de titre et auteur qui se trouvent à l'intérieur du document pour les ajouter sous forme de métadonnées dans les

¹³ Entretien avec l'interviewé 9, Gestionnaire des sinistres, Genève, 27 avril 2022.

champs correspondant. Selon le référent SIT¹⁴ interrogé, un ancien formulaire de saisie existe mais ne fait pas la part belle aux métadonnées, il n'est pas appliqué de manière systématique et les collaborateurs ne seraient pas encore suffisamment sensibilisés à ce propos. De son côté, la Spécialiste en gestion documentaire¹⁵, explique que dans le cadre d'un projet documentaire aux tpg nommé *Digidoc*, plusieurs personnes du métier ont participé, entre autres, à la création d'une nouvelle fiche descriptive. Le but est de simplifier la saisie pour le commun des collaborateurs qui alimentent eDocs. L'accent est mis sur la commodité et l'aspect pratique, il y a une volonté de réduire les métadonnées à l'essentiel et de structurer le contenu avec donc uniquement le titre, l'auteur et la date du document. Les utilisateurs sont invités à rédiger de manière exhaustive les titres et à éventuellement renseigner le sujet du document. eDocs permet la recherche plein texte des documents, ce qui encourage l'idée de ne pas encombrer les collaborateurs avec des exigences plus avancées pour la saisie des métadonnées.

Des métadonnées peuvent aussi être appliquées mais de manière plus diffuse dans ces deux autres applicatifs métier suivants : Pour SmartCockpit, l'outil de suivi des indicateurs de performance, il est possible d'ajouter une description à certains documents téléversés, mais cela est optionnel ; Pour Solu-Qiq, le manuel qualité de l'entreprise, il n'y a pas de pratique des métadonnées mais l'export d'un organigramme contient automatiquement son titre et une date qui correspond cependant au jour de l'export (constaté lors d'une démonstration du logiciel).

Concernant les autres applicatifs, aucune information de type métadonnée ne serait renseignée, les entretiens ont révélé qu'il n'y a pas de pratiques à ce sujet ni dans le cas d'un import, de l'utilisation au sein du logiciel ou lors d'un export.

Les documents des disques réseaux accessibles (P et Q) n'améliorent pas les résultats de cette évaluation, dans le meilleur des cas, un document Word ou Excel contient le titre, le nom de l'auteur, le nom de la personne ayant effectué le dernier enregistrement, le nom de l'entreprise (tpg) et l'horodatage. Toutes les métadonnées observées dans les disques réseaux sont enregistrées automatiquement. Il n'y a pas de métadonnées descriptives dans les documents ayant d'autres formats, en particulier les documents autres que la suite Microsoft Office. Les fichiers images, PDF ou DWG n'ont que des métadonnées techniques créées automatiquement. L'horodatage est présent dans tous les cas, mais il est possible de retrouver des incohérences comme des dates de modifications plus anciennes que les dates création. Les noms des auteurs ne sont pas uniformisés, il y a des noms complets suivis de la fonction métier et des noms simples avec prénom ou sans.

4.2.2 Recommandations pour les métadonnées

L'application des directives du tableau des métadonnées en annexe 6, supposera un réel changement d'habitudes pour les utilisateurs puisque les pratiques actuelles en matière de remplissage des métadonnées sont inexistantes dans la plupart des cas.

Comme vu plus haut (cf. 4.2.1 Analyse des métadonnées aux tpg), il n'est pas possible de présenter des informations complètes sur les métadonnées actuellement en usage dans les

¹⁴ Entretien avec l'interviewé 22, Responsable Infrastructures et Opérations, Genève, 11 avril 2022.

¹⁵ Entretien avec l'interviewé 23, Spécialiste en gestion documentaire des tpg, Genève, 25 avril 2022.

applicatifs métier, puisque les informations fournies lors des entretiens ne sont pas détaillées. Les entretiens ont aussi révélé des hésitations fréquentes à l'heure d'aborder la question des métadonnées. Pour dissiper les doutes, l'ouverture des propriétés des fichiers durant les démonstrations a permis de constater que dans la quasi-totalité des cas, il n'y a pas de métadonnées significatives pour renseigner les documents. Notons que certains logiciels sont utilisés par plusieurs DA de manière différente et avec des contenus différents. Les référent-e-s métiers interrogés n'ont donc pas la possibilité de donner des précisions supplémentaires sur l'utilisation d'autres documents listés comme archivables dans un domaine qui se trouve hors de leur compétence. Par exemple, un collaborateur ou une collaboratrice travaillant en Finance & gestion et étant un-e référent-e métier d'un logiciel contenant des documents exploités tant par Finance & gestion que par les Ressources humaines, ne pourra répondre qu'à propos des documents faisant partie de son domaine.

Vu le nombre important de dossiers dans les applicatifs métier et dans les disques réseau, il serait difficile de travailler les recommandations des métadonnées à ce niveau de précision. De plus, il n'est pas possible de consulter ou de se familiariser avec les particularités de tous les contenus visés par ce travail. C'est pourquoi les recommandations ci-dessous sont de nature simple et essentielle pour laisser suffisamment d'espace et de marge à leur adaptation.

Dans le tableau en annexe 6, quatre métadonnées de description facultatives ne sont pas reprises pour l'archivage, il s'agit de : l'éditeur du document, du contributeur du document, du sujet du document et de la description du document. Ces indications découlent directement du document des Archives de l'État de Genève (2011) pour le projet Gal@atae et n'ont pas subi d'adaptation. En partant du principe que les métadonnées existantes dans un document doivent être conservées, il serait envisageable de faire de même avec ces quatre métadonnées. Ces métadonnées ne sont peut-être pas nécessaires pour le versement via Gal@atae, mais elles présentent néanmoins des aspects intéressants pour compléter les entités responsables du document (contributeur et éditeur), ainsi que pour élargir les opportunités de recherche et de définition du document avec les champs sujet et description. Ces métadonnées sont d'ailleurs présentes parmi la liste des 15 éléments du Dublin Core leur reprise au moment de l'archivage devrait donc être à considérer.

Pour les applicatifs métier et les disques réseaux, l'application des directives du tableau avec les métadonnées minimales est à mettre à l'essai.

Les métadonnées déjà présentes dans les quatre applicatifs où le remplissage des informations est déjà protocolé (cf. 4.2.1 Analyse des métadonnées aux tpg), ne devraient pas subir d'altérations, et l'application du tableau serait complémentaire.

Pour les disques réseaux et les autres applicatifs métier où il n'y a pas de réelles pratiques à cet égard, l'imposition des métadonnées minimales telles qu'indiquées dans le tableau pourrait résulter une tâche difficile. Au tel cas, il faudrait s'assurer qu'un set minimal soit au moins remplis, en priorité le titre et la date dans les situations moins évidentes.

4.3 Création d'un protocole de versement

Le versement des documents physiques est un processus bien connu, bien que cette action ne disparaisse pas avec le numérique, la croissance exponentielle des documents numériques pousse les archivistes à prévoir les vagues de versements et à prioriser l'automatisation des processus. Selon Texier (2016) :

« L'accroissement de la production documentaire numérique oblige les organisations (publiques et privées) à repenser leurs procédures. Elles devront donc prendre en considération les aspects techniques et fonctionnels proposés par les systèmes d'archivage électronique proposés par les éditeurs. Et regarder de près l'interopérabilité entre les applications de production et les applications de consultation ainsi que les considérations techniques (création de métadonnées, techniques de versement...). »

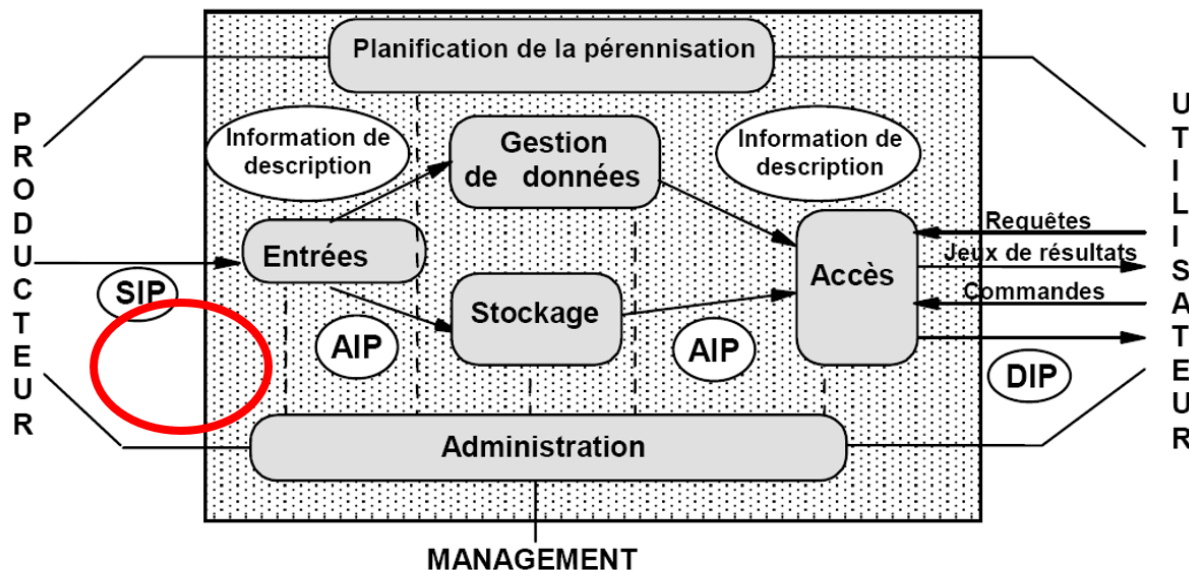
La fonction de versement est intégrée au sein du modèle OAIS qui aborde les fonctions de pérennisation d'archivage électronique. Selon ce modèle (Consultative Committee for Space Data System 2012 p. 1.16), le protocole de versement ou *Submission Agreement* en anglais, est un :

« accord négocié entre l'OAIS et le Producteur, spécifiant un modèle de données, et toute disposition nécessaire, pour l'Opération de versement. Ce modèle identifie le format, le contenu et la structure logique utilisés par le Producteur et leur mode de représentation pour chaque support de livraison ou lors d'une télétransmission. »

Le protocole de versement permet donc de formaliser les responsabilités entre les producteurs de documents et les Archives afin de mener à bien les modalités de versement vers la plateforme de pérennisation et d'archivage, en l'occurrence la solution docuteam cosmos dans le cas des tpg.

Si l'on reprend la figure du modèle OAIS ci-dessous, le protocole de versement se situerait dans le cercle rouge, entre le Producteur et les entités Entrée et Administration.

Figure 9 : Protocole de versement au sein d'OAIS



(Boucon 2008, p. 4)

Le protocole de versement fait ainsi le lien entre le producteur des documents à archiver et l'Administration en charge des « des fonctions et services relatifs à l'exploitation d'ensemble du système d'archivage » (Consultative Committee for Space Data System 2012, p. 4.2). Il définit le SIP (Paquet d'informations à verser ou Submission Information Package en anglais), qui contient les documents associés aux métadonnées de description et de pérennisation. Il

met ensuite en le lien le SIP avec l'entité Entrée chargée de le recevoir et d'en générer un AIP (Paquet d'informations archivé ou Archival Information Package) en vue du stockage.

Dans cette démarche, le protocole de versement pose les bases du contexte de production des documents à archiver, définit les exigences et modalités relatives au versement et répartit les rôles et responsabilités des différents acteurs.

Lors de la création d'un protocole de versement, il est donc nécessaire d'identifier tous les éléments nécessaires au bon déroulement de l'opération de versement afin d'éviter les obstacles et difficultés liées à une préparation insuffisante. La méthodologie PAIMAS (Producer-Archive Interface Methodology Abstract Standard), est dans ce cas utile pour « identifier les différentes phases dans le processus de transfert d'information entre producteur et archive » et « définir les objectifs de chaque phase, les actions associées et les sorties attendues » (Boucon 2008, p. 2).

PAIMAS (Consultative Committee for Space Data System 2004, pp. 2.3-2.4) est un guide qui contient quatre phases :

- Phase préliminaire ;
- Phase de définition formelle ;
- Phase de transfert ;
- Phase de validation.

La phase préliminaire contextualise l'envergure du projet, les personnes impliquées et la volumétrie des archives. Dans cette phase, l'intérêt est mis sur la faisabilité du projet et sur sa planification.

La phase de définition formelle, cherche à définir précisément le paquet d'informations à verser et les modalités de transfert tels que les outils à utiliser, les accès ou les fréquences de transfert vers la plateforme de pérennisation et d'archivage. Cette phase débouche dans un premier protocole de versement.

La phase de transfert, met en pratique le versement du paquet d'informations, les tests effectués à ce moment permettent de solutionner les problèmes pouvant surgir afin d'assurer le bon déroulement de ce processus.

La phase de validation est la dernière phase de PAIMAS, il s'agit de valider les différents points du processus de versement et de vérifier que les objets numériques versés sont conformes à ce qui est exigé dans le protocole.

Les deux premières phases de PAIMAS constituent donc un document de travail préalable et clé pour la rédaction d'un protocole de versement.

4.3.1 Méthodologie et propositions

Dans ce travail on dénombre une quinzaine d'applicatifs métier et plusieurs disques réseau, ce qui impliquerait la création d'un protocole de versement pour chaque outil. Dans un souci de faisabilité, il a donc été décidé de se prêter à l'exercice pour un nombre limité de trois outils maximums à savoir :

- La GED eDocs, vu son importance au sein de l'entreprise et le nombre de documents listés pour l'archivage ;

- Isilog, puisque ce logiciel va être décommissionné ;
- Le disque réseau Q, puisque presque tous les dossiers listés pour l'archivage sont accessibles contrairement aux autres disques réseau.

Afin de rédiger les protocoles de versement, une liste simplifiée des différentes phases PAIMAS est utilisée dans le cadre de ce travail. Il s'agit d'une liste reprenant chaque étape de la procédure de questionnement mise à disposition lors des cours académiques (Dubois 2019a). La liste est entièrement basée sur les recommandations PAIMAS et la numérotation des étapes correspond au document original (Consultative Committee for Space Data System 2004, pp. 2.3-2.4), ce qui permet de d'accéder à une description plus détaillée de chaque phase en cas de besoin. Les deux premières phases de cette liste sont disponibles en annexe 7, les étapes sont complémentées de prises de notes en bleu.

Lors du renseignement de la liste simplifiée PAIMAS, un premier constat s'impose : beaucoup plus d'informations sont nécessaires pour répondre précisément aux divers points listés. Cela implique plus d'échanges avec les producteurs qui permettent de concrétiser les démarches et avec qui la collaboration doit être plus étroite.

En effet, les informations récoltées lors des entretiens auprès des référents métier et SIT ne sont malheureusement pas suffisantes pour répondre à chacun des points de la phase préliminaire et il n'est donc pas non plus possible d'apporter de précisions à la phase de définition formelle.

À ce stade, les trois protocoles de versement proposés en annexe 8, 9 et 10 ne constituent donc qu'une ébauche conceptuelle. Ils ont vocation à donner une idée de ce à quoi peut ressembler un tel document pour eDocs, Isilog ou les dossiers du disque Q.

Ces protocoles de versement ont été créés sur la base d'un exemple des Archives d'État du Valais pour les archives électroniques et émanant du cursus académique (Dubois 2019b). Ce document exemple contient six chapitres pour couvrir les points essentiels de l'accord :

1. Le Contexte : permet d'introduire les informations à archiver, les enjeux et la raison d'être du protocole de versement en question ;
2. L'Information sur le producteur : pour indiquer l'origine des documents versés ;
3. La Description des documents versés : développe les informations sur la typologie des documents, leur période de création, la volumétrie et les droits et moyens d'accès aux documents une fois déposés ;
4. Le Versement : schématise l'arborescence des documents, les formats acceptés, les modalités du nommage des fichiers, les exigences archivistiques et le processus du versement ;
5. Les Rôles et responsabilités : permet de répartir les tâches parmi les différentes parties prenantes au projet de versement ;
6. La Révision : permet d'indiquer la fréquence ou les raisons qui peuvent pousser à réviser le présent protocole de versement.

Plusieurs points restent à préciser dans les trois protocoles en annexe.

Dans les protocoles de versement en annexe, il est proposé un nommage systématique incluant un numéro d'identifiant unique, la date et l'intitulé du document. Les métadonnées doivent correspondre aux recommandations établies, de-même pour les formats qui dans

certains cas devront être convertis selon les recommandations de ce travail. Les processus de versement prévus sont similaires dans tous les cas, il est proposé de verser les fichiers extraits des applicatifs métier dans le dossier de transfert existant actuellement dans le disque réseau Q et d'opérer dans ce dossier les adaptations en vue du versement sur docuteam cosmos.

La volumétrie des documents de la GED eDocs (annexe 8) à archiver n'est pas connue. Il s'agit d'estimer la volumétrie des 181 documents et séries de documents listés en annexe 2. Le numéro d'identifiant de chaque fichier devrait être annoncé afin de pouvoir les retrouver au sein de la GED pour quantifier leur volumétrie totale. L'accès à ces documents permettrait aussi de renseigner les dates extrêmes.

En dehors des référents métier et SIT de la GED, il est proposé d'intégrer les répondants documentaires de chaque domaine d'activité. Les documents sur eDocs découlent du travail de tous les domaines d'activité des tpg. Les répondants documentaires pourraient donc être mobilisés afin d'identifier précisément les documents listés en annexe 3.

Pour Isilog (annexe 9), les zones floues sont similaires, l'information de volumétrie est vague et les dates extrêmes sont déduites par l'ancienneté du logiciel introduit en 2008. Les documents listés pour Isilog proviennent du Conseil administratif (CA) et de la Direction générale (DG), il s'agit surtout de suivi d'audits qui peuvent comporter des pièces-jointes. Les informations sous la forme d'extraction des audits et leur organisation ne ressort pas de manière évidente lors des entretiens, raison pour laquelle les pratiques de versement à partir d'Isilog sont à revoir en détail.

Enfin, concernant le disque réseau Q (annexe 10), un des quatre dossiers à archiver est restreint d'accès. Il existe aussi un cas présentant une arborescence des dossiers complexe, le dossier contenant les plans des véhicules est sous-divisé en multiples sous-dossiers et comporte de nombreux formats différents. Les services producteurs des documents sont connus mais contrairement aux applicatifs métier, il n'y a pas de référents. Les services SIT sont tout de même inclus en fin de document afin d'apporter en cas de besoin un soutien pour les points les plus techniques.

5. Conclusion

Ce travail mandaté par les Archives des tpg, est une occasion pour entamer les premières réflexions en matière d'archivage électronique au sein de cette entreprise enracinée dans l'histoire de la mobilité à Genève. Les recommandations formulées à travers de cet exercice, ont pour objectif de poser des bases pour la prise d'actions concrètes en vue du transfert des informations archivables vers la plateforme de pérennisation et d'archivage.

L'organisation, la conduite et l'analyse des entretiens semi-directifs menés auprès des responsables des applicatifs métier de l'entreprise, ont constitué les principales tâches de ce travail. Les résultats de l'état des lieux permettent de constater que si la gestion courante des documents numériques est plutôt fluide, les questions de conservation des documents et d'extraction sont moins bien connues, elles sont encore peu prises en compte dans les réflexions. Cela a parfois engendré des retours timides avec des informations souvent peu détaillées sur la manière dont les données sont conservées ou extractibles.

Malgré ces lacunes, ce travail s'efforce d'émettre des propositions en s'appuyant également sur des bases théoriques et nourrit l'espoir de servir de d'appui à la réalisation des projets à venir.

Enfin, les technologies sont amenées à évoluer, des nouvelles montées de version logicielles sont à prévoir aux tpg dans le court et moyen terme. Dans cette perspective, il conviendrait d'inclure la question de la gestion archivistique des documents et dossiers, dès l'introduction d'une l'application ou de l'implémentation d'une nouvelle version.

Bibliographie

ARCHIVES D'ÉTAT DE GENÈVE, 2011a. Formats de fichiers adaptés à l'archivage électronique à moyen et long terme. *Archives d'Etat de Genève (ge.ch)* [en ligne]. 19 novembre 2011. [Consulté le 14 juin 2022]. Disponible à l'adresse : https://ge.ch/archives/media/site_archives/files/imce/pdf/procedures/3_20111019_formats_archivage_etatge_v.pdf

ARCHIVES D'ÉTAT DE GENÈVE, 2011b. Liste des métadonnées à implémenter dans une application GED. *Archives d'Etat de Genève (ge.ch)* [en ligne]. 9 juin 2011. [Consulté le 22 juin 2022]. Disponible à l'adresse : https://ge.ch/archives/media/site_archives/files/imce/pdf/procedures/1_20111019_metadonnees_obligatoire_etatge_v.pdf

ARCHIVES DE L'ÉTAT DE BERNE, 2022. Qu'est-ce que l'archivage numérique à long terme? *Archives de l'État de Berne* [en ligne]. [Consulté le 13 juin 2022]. Disponible à l'adresse : <https://www.staatsarchiv.sta.be.ch/fr/start/fuer-behoerden/was-ist-digitale-langzeitarchivierung.html>

ARCHIVES FÉDÉRALES SUISSES (AFS), 2020. Normes et standards pour l'archivage de documents numériques – Formats de fichiers adaptés à l'archivage. *Archives fédérales suisses* [en ligne]. Avril 2020. [Consulté le 14 juin 2022]. Disponible à l'adresse : https://www.bar.admin.ch/dam/bar/fr/dokumente/konzepte_und_weisungen/archivtaugliche-dateiformate.1.pdf.download.pdf/formats_de_fichiersadaptesalarchivage.pdf

ARCHIVES NEW ZEALAND, 2022a. Metadata for information and records. *Archives New Zealand* [en ligne]. 22 juin 2022. [Consulté le 22 juin 2022]. Disponible à l'adresse : <https://www.archives.govt.nz/manage-information/how-to-manage-your-information/implementation/metadata/metadata-for-information-and-records>

ARCHIVES NEW ZEALAND, 2022b. Requirements for metadata. *Archives New Zealand* [en ligne]. 22 juin 2022. [Consulté le 22 juin 2022]. Disponible à l'adresse : <https://www.archives.govt.nz/manage-information/how-to-manage-your-information/implementation/metadata/requirements-for-metadata>

Aspirateur de site web. *Wikipédia : l'encyclopédie libre* [en ligne]. Dernière modification de la page le 5 novembre 2019 à 12:55. [Consulté le 2 août 2022]. Disponible à l'adresse : http://fr.wikipedia.org/w/index.php?title=Aspirateur_de_site_web&oldid=164208708

AUSTRALASIAN DIGITAL RECORDKEEPING INITIATIVE (ADRI), 2020. Sustainable Digital Formats for Creating and Using Records. *Council of Australasian Archives and Records Authorities* [en ligne]. Avril 2020. [Consulté le 14 juin 2022]. Disponible à l'adresse : <https://www.caara.org.au/wp-content/uploads/2020/07/Sustainable-Digital-File-Formats-for-Creating-and-Using-Records-V1.0-April-2020.pdf>

BIBLIOTHÈQUE ET ARCHIVES CANADA (BAC), 2015. Lignes directrices sur les formats de fichier à utiliser pour transférer des ressources documentaires. *Bibliothèque et Archives Canada* [en ligne]. 6 février 2015. [Consulté le 15 juin 2022]. Disponible à l'adresse : <https://www.bac-lac.gc.ca/fra/services/gestion-ressources-documentaires-gouvernement/lignes-directrices/Pages/lignes-directrices-formats-fichier-transferers-ressources-documentaires.aspx>

BIBLIOTHÈQUE NATIONALE DE FRANCE (BNF), 2008. Guide d'utilisation du Dublin Core (DC) à la BnF : Dublin Core simple et Dublin Core qualifié, avec indications pour utiliser le profil d'application de TEL. *Bibliothèque nationale de France* [en ligne]. 8 septembre 2008.

Version 2. [Consulté le 23 juin 2022]. Disponible à l'adresse : https://www.bnf.fr/sites/default/files/2019-02/guide_dublin_core_bnf_2008.pdf

BIBLIOTHÈQUE NATIONALE DE FRANCE (BNF), 2021. Formats de données pour la préservation à long terme : la politique de la BnF. *Bibliothèque nationale de France* [en ligne]. 9 avril 2021. Version 2. [Consulté le 15 juin 2022]. Disponible à l'adresse : https://www.bnf.fr/sites/default/files/2021-04/politiqueFormatsDePreservationBNF_20210408.pdf

BIBLIOTHÈQUE NATIONALE DE FRANCE (BNF), 2022a. PREMIS : PREservation Metadata Implementation Strategies. *Bibliothèque nationale de France* [en ligne]. [Consulté le 24 juin 2022]. Disponible à l'adresse : <https://www.bnf.fr/fr/premis-preservation-metadata-implementation-strategies>

BIBLIOTHÈQUE NATIONALE DE FRANCE (BNF), 2022b. EAD : Encoded Archival Description. *Bibliothèque nationale de France* [en ligne]. [Consulté le 28 juin 2022]. Disponible à l'adresse : <https://www.bnf.fr/fr/ead-encoded-archival-description>

BIBLIOTHÈQUE NATIONALE SUISSE, 2022. Collections numériques BN et Archives Web Suisse. *Bibliothèque nationale suisse* [en ligne]. [Consulté le 29 juillet 2022]. Disponible à l'adresse : <https://www.e-helvetica.nb.admin.ch/>

BOUCON, Danièle, 2008. PAIMAS et PAIS. *Association ARISTOTE* [en ligne]. 19 mai 2008. [Consulté le 5 juillet 2022]. Disponible à l'adresse : <http://pin.association-aristote.fr/lib/exe/fetch.php/public/presentations/2008/pin20080519-daniele-boucon-paimas-pais.pdf>

CAPLAN, Priscilla, 2009. Comprendre PREMIS. *Library of Congress* [en ligne]. 1 février 2009. [Consulté le 24 juin 2022]. Disponible à l'adresse : http://www.loc.gov/standards/premis/Understanding-PREMIS_french.pdf

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021a. PDF/A-1. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 20 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/pdf-a-1_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021b. XLS. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 20 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/xls_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021c. XLSX. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 20 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/xlsx_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021d. CSV. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 20 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/csv_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021e. DWG. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 20 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/dwg_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021f. DXF. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 20 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/dxf_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021g. JPEG. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 21 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/jpeg_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021h. PNG. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 21 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/png_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021i. HTML. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 21 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/HTML_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021j. SQL script. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 21 juin 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/sql-script_fr.html?highlight=sql

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2021k. Introduction. *KOST-CECO* [en ligne]. Décembre 2021. Version 6.2. [Consulté le 3 août 2022]. Disponible à l'adresse : https://kost-ceco.ch/cms/kad_introduction_fr.html

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2022a. *KOST-CECO* [en ligne]. [Consulté le 14 juin 2022]. Disponible à l'adresse : <https://kost-ceco.ch/cms/bienvenue.html>

CENTRE DE COORDINATION POUR L'ARCHIVAGE À LONG TERME DE DOCUMENTS ÉLECTRONIQUES (CECO), 2022b. Recherche impossible dans le PDF. *KOST-CECO* [en ligne]. [Consulté le 20 juin 2022]. Disponible à l'adresse : <https://kost-ceco.ch/cms/recherche-impossible-dans-le-pdf.html>

COMITÉ CONSULTATIF POUR LES SYSTÈMES DE DONNÉES SPATIALES (CCSDS), 2012. Modèle de référence pour un système ouvert d'archivage d'information (OAIS). *CCSDS* [en ligne]. Juin 2012. Livre Magenta. CCSDS 650.0-M-2 (F). [Consulté le 13 juin 2022]. Disponible à l'adresse : <https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf>

CONSEIL INTERNATIONAL DES ARCHIVES, 2000. ISAD(G) : norme générale et internationale de description archivistique. Conseil international des archives [en ligne]. [Consulté le 29 juin 2019]. Disponible à l'adresse : https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_FR.pdf

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS), 2004. Producer-Archive Interface Methodology Abstract Standard (PAIMAS). *CCSDS* [en ligne]. Mai 2004. Livre Magenta. CCSDS 651.0-M-1. [Consulté le 5 juillet 2022]. Disponible à l'adresse : <https://public.ccsds.org/Pubs/651x0m1.pdf>

COUTAZ, Gilbert, 2016. La croissance et la maîtrise des masses documentaires. *arbido* [en ligne]. 2016. n°3. [Consulté le 7 juillet 2022]. Disponible à l'adresse : <https://arbido.ch/fr/edition-article/2016/d%C3%A9truire-pour-conserver/la-croissance-et-la-ma%C3%A9trise-des-masses-documentaires>

CRIDER, Michael, 2017. What Is a .DOCX File, and How Is It Different from a .DOC File in Microsoft Word? *How-To Geek* [en ligne]. 2 mai 2017. [Consulté le 3 août 2022]. Disponible à l'adresse : <https://www.howtogeek.com/304622/what-is-a-docx-file-and-how-is-it-different-from-a-doc-file-in-microsoft-word/>

DCMI USAGE BOARD, 2012. Dublin Core™ Metadata Element Set. *Dublin Core* [en ligne]. 14 juin 2012. Version 1.1. [Consulté le 23 juin 2022]. Disponible à l'adresse : <https://www.dublincore.org/specifications/dublin-core/dces/>

DUBOIS, Alain, 2019a. *Méthodologie PAIMAS – Liste simplifiée des différentes phases* [document Word]. Support de cours : Cours « Gestion électronique et numérisation des archives », Haute école de gestion de Genève, filière Information documentaire, année académique 2018-2019.

DUBOIS, Alain, 2019b. *Protocole de versement d'archives électroniques : Archivage des actes législatifs du canton du Valais* [document Word]. Support de cours : Cours « Gestion électronique et numérisation des archives », Haute école de gestion de Genève, filière Information documentaire, année académique 2018-2019.

DUNANT GONZENBACH, Anouk, 2013. La gouvernance des documents électroniques dans l'administration cantonale genevoise : genèse et mise en œuvre. *Ressi* [en ligne]. 18 décembre 2013. n°14. [Consulté le 13 juin 2022]. Disponible à l'adresse : http://www.ressi.ch/num14/article_93

DURANTI, Luciana, 2021. The Concept of Digital Record according to InterPARES. *arbido* [en ligne]. 2021. n°4. [Consulté le 13 juin 2022]. Disponible à l'adresse : <https://arbido.ch/fr/edition-article/2021/das-dokument/concept-of-digital-record-interpares>

FILEINFO, 2021. IDW File Extension. *FileInfo* [en ligne]. 22 juillet 2021. [Consulté le 20 juin 2022]. Disponible à l'adresse : <https://fileinfo.com/extension/idw>

GOT [Pseudonyme], 2010. Vous avez dit « format pérenne » ? *Les petites cases* [en ligne]. 5 mai 2010, 19:23. [Consulté le 13 juin 2022]. Disponible à l'adresse : <http://www.lespetitescases.net/vous-avez-dit-format-perenne>

GOT [Pseudonyme], 2018. La conservation des données numériques n'est ni une révolution, ni inatteignable. *Les petites cases* [en ligne]. 15 janvier 2018, 10:44. [Consulté le 13 juin 2022]. Disponible à l'adresse : <http://www.lespetitescases.net/la-conservation-des-donnees-numeriques-n-est-ni-une-revolution-ni-inatteignable>

GROUPE AGS, 2021. 8 bonnes raisons d'améliorer votre gestion documentaire. *AGS-Archivage* [en ligne]. 15 mars 2021. [Consulté le 7 juillet 2022]. Disponible à l'adresse : <https://www.ags-archivage.com/news/8-bonnes-raisons-d-ameliorer-votre-gestion-documentaire/>

NATIONAL ARCHIVES OF AUSTRALIA, 2022. Minimum metadata set. *National Archives of Australia* [en ligne]. [Consulté le 22 juin 2022]. Disponible à l'adresse : <https://www.naa.gov.au/information-management/information-management-standards/australian-government-recordkeeping-metadata-standard/minimum-metadata-set>

LIBRARY OF CONGRESS, 2020. Microsoft Outlook Item (MSG). *Library of Congress* [en ligne]. 26 février 2020. [Consulté le 21 juin 2022]. Disponible à l'adresse : <https://www.loc.gov/preservation/digital/formats/fdd/fdd000379.shtml>

LIBRARY OF CONGRESS, 2021. Visio VSDX Drawing File Format. *Library of Congress* [en ligne]. 10 août 2021. [Consulté le 20 juin 2022]. Disponible à l'adresse : <https://www.loc.gov/preservation/digital/formats/fdd/fdd000021.shtml>

LIBRARY OF CONGRESS, 2022a. Sustainability of Digital Formats: Planning for Library of Congress Collections. *Library of Congress* [en ligne]. 26 mai 2022. [Consulté le 15 juin 2022]. Disponible à l'adresse : <https://www.loc.gov/preservation/digital/formats/index.html>

LIBRARY OF CONGRESS, 2022b. Encoded Archival Description. *Library of Congress* [en ligne]. 29 mars 2022. [Consulté le 28 juin 2022]. Disponible à l'adresse : <https://www.loc.gov/ead/>

PREIS EDITORIAL COMMITTEE, 2015. PREMIS Data Dictionary for Preservation Metadata. *Library of Congress* [en ligne]. Juin 2015. Version 3.0. [Consulté le 24 juin 2022]. Disponible à l'adresse : <http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>

QUEENSLAND GOVERNEMENT, 2021. File formats for long-term digital records. *Queensland Government* [en ligne]. 5 mai 2021. [Consulté le 13 juin 2022]. Disponible à l'adresse : <https://www.forgov.qld.gov.au/information-and-communication-technology/recordkeeping-and-information-management/recordkeeping/store-protect-and-care-for-records/store-protect-and-care-for-digital-records/file-formats-for-long-term-digital-records>

RIETSCH, Jean-Marc, et al., 2010. *Mise en œuvre de la dématérialisation : cas pratiques pour l'archivage électronique*. Paris : Dunod. ISBN 978-2-10-055688-5

SCHAERER, Aurélie, 2021. *Gouvernance documentaire : Règlement* [fichier PDF]. Document interne à l'entreprise.

SCHAERER, Aurélie, SCHNEIDER, Cynthia et THONNEY, Delphine, 2022. *Archives : Jeu de métadonnées obligatoires* [fichier Microsoft Word]. Document interne à l'entreprise.

SCORM, 2022. SCORM [en ligne]. [Consulté le 3 août 2022]. Disponible à l'adresse : <https://scorm.com/>

SOCIETY OF AMERICAN ARCHIVISTS, 2004. Description Archivistique Encodée : Dictionnaire des balises. *France Archives* [en ligne]. Octobre 2004. Version 2002. [Consulté le 28 juin 2022]. Disponible à l'adresse : https://francearchives.fr/file/0def64f5a10f3f1ae03fdea59399a3e0755ef157/static_1066.pdf

TEXIER, Bruno, 2016. Le versement d'archives à l'ère numérique. *Archimag* [en ligne]. 26 février 2016. [Consulté le 5 juillet 2022]. Disponible à l'adresse : <https://www.archimag.com/archives-patrimoine/2016/02/26/versement-archives-ere-numerique>

THE NATIONAL ARCHIVES, 2022. The technical registry PRONOM : About. *The National Archives* [en ligne]. [Consulté le 15 juin 2022]. Disponible à l'adresse : <https://www.nationalarchives.gov.uk/aboutapps/pronom/default.htm>

THONNEY, Delphine, 2020. *Élaboration d'une stratégie de valorisation des archives des TPG* [en ligne]. Genève : Haute école de gestion de Genève. Travail de Master. [Consulté le 17 juin 2022]. Disponible à l'adresse : <https://sonar.ch/hesso/documents/314918>

TRANSPORTS PUBLICS GENEVOIS, 2020. *Politique de gouvernance documentaire* [fichier PDF]. 9 septembre 2020. Document interne à l'entreprise

TRANSPORTS PUBLICS GENEVOIS, 2021a. Rapport annuel de gestion 2020. *Transports publics genevois* [en ligne]. [Consulté le 16 juin 2022]. Disponible à l'adresse : https://www.tpg.ch/sites/default/files/2021-07/TPG-RAG2020_WEB_pageparpage_FINAL.pdf

TRANSPORTS PUBLICS GENEVOIS, 2022a. Rapport annuel de gestion 2021. *Transports publics genevois* [en ligne]. [Consulté le 16 juin 2022]. Disponible à l'adresse : https://www.tpg.ch/sites/default/files/2022-05/FINAL_TPG_RAPPORT_ANNUEL_2021_WEB.pdf

TRANSPORTS PUBLICS GENEVOIS, 2022b. Cap2030, Mobiles ensemble. *Transports publics genevois* [en ligne]. [Consulté le 7 juillet 2022]. Disponible à l'adresse : <https://www.tpg.ch/fr/cap2030>

TRANSPORTS PUBLICS GENEVOIS, 2022c. *Référentiel services et traitements* [fichier PDF]. 31 janvier 2022. Document interne à l'entreprise

TRANSPORTS PUBLICS GENEVOIS, 2022d. Archives des transports publics genevois. *Transports publics genevois* [en ligne]. [Consulté le 27 juillet 2022]. Disponible à l'adresse : <https://www.tpg.ch/fr/nous-connaitre/acteur-majeur-de-la-mobilite/archives-des-transports-publics-genevois>

Wiki. *Wikipédia : l'encyclopédie libre* [en ligne]. Dernière modification de la page le 6 juin 2022 à 20:51. [Consulté le 16 juin 2022]. Disponible à l'adresse : <https://en.wikipedia.org/w/index.php?title=Wiki&oldid=1091861967>

Wikitexte. *Wikipédia : l'encyclopédie libre* [en ligne]. Dernière modification de la page le 9 juin 2022 à 14:43. [Consulté le 16 juin 2022]. Disponible à l'adresse : <http://fr.wikipedia.org/w/index.php?title=Wikitexte&oldid=194398618>

Annexe 1 : Diagramme de Gantt

	février		mars					avril					mai				juin				juillet				août	
Tâches	14.02-20.02	21.02-27.02	28.02-06.03	07.03-13.03	14.03-20.03	21.03-27.03	28.03-03.04	04.04-10.04	11.04-17.04	18.04-24.04	25.04-01.05	02.05-08.05	09.05-15.05	16.05-22.05	23.05-29.05	30.05-05.06	06.06-12.06	13.06-19.06	20.06-26.06	27.06-03.07	04.07-10.07	11.07-17.07	18.07-24.07	25.07-31.07	01.08-07.08	08.08-14.08
■ Initialisation																										
• Prise de contact avec mandante tpg																										
• Prise de contact avec directeur TM																										
• Remise du protocole d'accord signé à Responsable IS																										
• Rédaction du cahier des charges (remise v1 le 11.03)																										
• Revue de littérature																										
■ État des lieux																										
• Inventaire des applicatifs métier																										
• Détail de l'usage et fonctionnalités des app. métier																										
• Entretiens avec les responsables métier et applicatif, analyse des pratiques et identification du processus de gestion des documents à archiver																										
• Création d'un tableau synoptique des app. métier																										
■ Recommandations de conservation																										
• Recommandations sur les formats pérenns																										
• Recommandation sur les métadonnées minimales																										
• Propositions pour un protocole versement																										
■ Travail de Master																										
• Rédaction																										
• Relecture																										
• Remise du travail (au plus tard 15.08)																										

Annexe 2 : Grille de lecture

Titre	Année	Sujet et type de source	Synthèse/Commentaires	Notions générales	Normes / modèles	Versement	Formats	Métadonnées	Présentation de cas
PAIMAS	2004	PDF	Producer-Archive Interface Methodology Abstract Standard		x	x			
PAIMAS et PAIS (CNES)	2008	PDF	Présentation de PAIMAS par le groupe PIN.		x	x			
Comprendre PREMIS	2009	Guide pour comprendre PREMIS (en français), PDF	Ce guide à pour vocation de donner une idée de ce qu'est PREMIS et de servir d'introduction.		x			x	
L'archivage numérique à long terme: les débuts de la maturité ?	2009	Livre	Aborde la pérennité de l'archivage numérique	x			x		
Politique d'archivage à long terme des documents électroniques aux AEG	2010	AEG, PDF		x	x	x			
Mise en oeuvre de la dématérialisation	2010	Livre (aussi sur ScholarVox)	Enjeux des Archives orienté entreprise, étapes vers la mise en service d'archivage électronique, "cas pratiques pour l'archivage électronique"	x					
Vous avez dit « format pérenne » ?	2010	Billet de blog	Commentaire critique sur la notion de "format pérenne".				x		
Formats de fichiers adaptés à l'archivage électronique à moyen et long terme	2011	PDF	Guide de recommandations pour les formats à utiliser à moyen et long terme pour l'administration et les AEG.				x		

Liste des métadonnées à implémenter dans une application GED	2011	AEG, PDF	Liste des métadonnées obligatoires pour un système GED aux AEG.					x	
Les chantiers du numérique	2012	Livre (aussi sur ScholarVox)	"dématisation des archives et métiers de l'archiviste"	x	x				
Modèle de référence OAIS	2012	PDF	Version 2012 traduite en français.		x	x		x	
Présentation des normes existantes en matière de gestion des documents électroniques	2013	Archives d'État de Genève, PDF	Synthèse des normes existantes		x				
La gouvernance des documents électroniques dans l'administration cantonale genevoise : genèse et mise en œuvre	2013	Ressi	Article qui expose la problématique et les principales étapes pour une meilleure gestion électronique des documents à l'État de Genève	x			x	x	x
eCH-0164 Modèle de cycle de vie d'une affaire (processus, dossiers et documents)	2014	Fournit des éléments permettant de décrire de manière générique des modèles de cycles de vie, eCH	Présente un schéma du cycle de vie au point 6.2.		x				
eCH-0160 Interface de versements archivistiques	2015	Préparation d'un versement numérique, spécification du paquet SIP, eCH	"spécifie une interface de versement des dossiers et des documents qui proviennent de systèmes de gestion des dossiers, d'applications spécialisées/bases de données et de collections de fichiers à une archive"		x	x			

PREMIS version 3.0	2015	Dernière version de PREMIS, PDF	"The PREMIS Data Dictionary for Preservation Metadata is the international standard for metadata to support the preservation of digital objects and ensure their long-term usability". Le PDF est accessible via le site : www.loc.gov/standards/premis/		x			x	
L'archive dans quinze ans	2015	Livre (aussi sur ScholarVox)	Projection avenir, "nouveaux fondements"	x					
Archivistes de 2030	2015	Livre (aussi sur ScholarVox)	Projection avenir	x					
La croissance et la maîtrise des masses documentaires	2016	arbido	Aborde la question du poids du numérique	x					
Le versement d'archives à l'ère numérique	2016	Problématiques du versement numérique, Archimag	Automatisation du versement, différences du versement numérique			x			
Archivage des dossiers médicaux électroniques dans le Canton de Genève	2018	mise en place de canaux de transferts pour archiver les documents, arbido	Problématique du passage du versement analogique au numérique aux HUG et solution	x					x
Digital Processing Framework	2018	PDF	Description des tâches pour créer un SIP/AIP/DIP, archiver métadonnées techniques etc.			x		x	
Les 14 normes et référentiels de la dématérialisation à connaître absolument	2020	Archimag	Synthèse des normes existantes		x				
Les archives électroniques (Deuxième édition)	2020	Livre	Petit guide proposé par l'AAF	x					
Archivage électronique : il faut voir au-delà de la simple conservation	2020	Article en ligne, archivesfacotry.com		x					

Normes et standards pour l'archivage de documents numériques - Formats de fichier adaptés à l'archivage	2020	Archives fédérales suisses, PDF	Document qui offre un aperçu des formats actuellement acceptés par les AFS.				x		
The Concept of Digital Record according to InterPARES	2021	arbido	Problématique de la préservation à long terme et de l'authenticité	x			x		
Guide de gestion des documents	2021	Concept général de l'archivage électronique, PDF et site en ligne	Guide pour administration cantonale du Valais. Aborde certains points sur les archives électroniques.	x					
Quelles sont les tendances de l'archivage électronique en 2021 ?	2021	Archimag	"la problématique de la préservation des documents signés et celle d'un mode d'exploitation dédié ou mutualisé du SAE apparaissent comme deux tendances majeures pour 2021."	x					
Marche à suivre pour le versement de paquets d'archives électroniques aux Archives d'Etat de Genève	2021	PDF	Marche à suivre pour le versement d'archives électroniques pour les AEG.			x			
Formats de données pour la préservation à long terme : la politique de la BnF	2021	PDF	"Ce document décrit la politique globale de la BnF vis-à-vis des formats de données dans la perspective de la préservation à long terme de l'information".				x		
CECO	2022	Centre de coordination pour l'archivage à long terme de documents électroniques, site internet du CECO	Évaluations et recommandations sur les formats d'archivage				x		

Dématique	2014-2022	Série de plusieurs billets de blog (+30 billets), Convergence (AAQ)	Thèmes : GED vs SAE, Normes et standards, fonction versement et transferts, paquet d'informations	x	x	x	x	x	
PREMIS : PREservation Metadata Implementation Strategies	Page web	Présentation du dictionnaire de données PREMIS, Page web de la BNF			x			x	
Package Handler (AFS)	Page web	Page web des AFS	"Avec le logiciel Package Handler développé par les Archives fédérales, les paquets numériques peuvent être créés, édités, validés et consultés"			x			
File formats for long-term digital records	Page web	Queensland Government	Recommandations sur les formats				x		
Sustainability of Digital Formats: Planning for Library of Congress Collections	Page web	Ressources et informations à propos des formats	Qu'est-ce qu'un format? Relations entre les formats... Exemples de sept indices permettant d'évaluer si un format est durable.				x		
PRONOM	Page web	Registre d'informations techniques sur les formats mise en ligne par les Archives nationales du Royaume-Uni.	Permet de rechercher un format pour obtenir des informations techniques à propos et s'il est prioritaire ou non par rapport à un autre.				x		
Lignes directrices sur les formats de fichier à utiliser pour transférer des ressources documentaires	Page web	Page web	Liste des formats acceptés par la BAC.				x		
Dublin Core	Site web	Dublin Core (Site web)	Propose 15 éléments de métadonnées minimales pour la description (Dublin Core 15 elements).		x			x	

Archivematica	Site web	Site web, outil	"Archivematica is a web- and standards-based, open-source application which allows your institution to preserve long-term access to trustworthy, authentic and reliable digital content"	x					
Metadata for information and records	Site web	Archives New Zealand (Site web)	Résumé sur la métadonnée dans les archives					x	
Requirements for metadata	Site web	Archives New Zealand (Site web)	Résumé sur les métadonnées minimales					x	
Minimum metadata set	Site web	National Archives of Australia (Site web)						x	

Annexe 3 : Liste des archives à valeur historique et leur emplacement

N°	Applicatif	Descriptif	Séries de documents / dossiers	DUL	Sort final	Réf. Règle cdc	Remarques
1	360Learning		Formations tpg, support de formation, de cours	5	TPG	7.28 E, 7.37 RH	
2	Allegro		APOP (Analyse des prestations et objectifs professionnels)	10	ECH	7.38 RH	Conservation de toutes les APOP des personnes dont le dossier RH est échantillonné (cadre. Cadre supérieur, directeur et nom de famille commençant par lettre B)
3	AtoM		Inventaires	5	TPG	4.13 FG	Pas besoin de faire une extraction d'AtoM
6	Configuration management database (CMDB)			10	TPG	11.04 SIT	
7	Confluence		Cahier des charges de la ligne	10	TPG	12.19 E	Egalement dans Teams?
			Dossier création d'étape	10	TPG	12.18.1 E	
			Dossier mise à jour production étape	10	TPG	12.18.2 E	
			Comité d'architecture technique (Carch), Comité de changement informatique)	10	TPG	8.06 SIT, 8.07 SIT	
			Comité digital	10	TPG	8.09.1 SIT	
			Portefeuille de projet d'entreprise	5	TPG	1.10.01 DG	GED et confluence
9	Dir'AJ		Contrats	10	TPG	5.02 F&G	Le document scanné est dans la GED. Le contrat de prestations,

					le contrat de coopération Unireso, la Convention du Léman Passe font partie de cette série.
	Protection et accès documents, demandes de protection des données (LIPAD, RGPD, LPD)	10	TPG	4.05 DG	
	Avis de droit, interprétation de dispositions légales	10	TPG	5.01 DG	
	Dossier juridique - commentaire de contrat	15	TRI	5.17 DG	
	Pièces officielles (concession, actes notariés, licences, homologation, etc.)	10 / 20	TPG		Dénomination qui n'existe plus depuis 2022.
	Dossier lanceur d'alerte identifié	10	TPG	5.20.1 DG	
	Propriété intellectuelle, dépôt de marques	10	TPG	5.44 DG	
	Polices d'assurance	30	TPG	5.22 DG	
	Contrats de financement	10	TPG	5.01 FG	
	Séance OETHand	10	TPG	8.06 DG	Aussi dans la GED
	Concession pour des transports réguliers de voyageurs effectués à titre professionnel au moyen de tramways	10	TPG	12.33 MVC	Aussi dans la GED
	Autorisation cantonale pour des transports réguliers de voyageurs effectués à titre professionnel au moyen de bus	10	TPG	12.34 MVC	Aussi dans la GED
	Concession de lignes pour les autobus	10	TPG	12.36 MVC	Aussi dans la GED
	Concession pour les trolleybus	10	TPG	12.37 MVC	Aussi dans la GED
	Convention d'objectifs LEN (loi sur l'énergie)	Vie bâtiment	TPG	2.05 IB	
	Dossier de recherche historique	10	TPG	4.11.1 FG	
	Dossier de valorisation archivistique	10	TPG	4.11.2 FG	

10	GestE		Conventions financières	10	TPG	5.01 FG	
			Dossiers contentieux extra-judiciaire (avec ou sans blessé)	10 ou 20	ECH	5.31 DG, 5.31.1 DG	10% des dossiers à échantillonner
			Dossiers contentieux procédures judiciaires	10	TPG	5.30 DG	
			Dossier du chantier	10	TRI	12.15 E	Tri sur la base de l'importance du chantier.
			Dossier manifestation annuelle	5	ECH	12.16 E	Critère d'échantillonnage : conserver tous les dossiers manifestations des années terminant par "0" : ...1990, 2000, 2010, ...
			Dossier de manifestation exceptionnelle	15	TRI	12.17 E	Tri = Conservation de toutes les manifestations uniques ou dont la récurrence est supérieure à 4 ans. Les autres manifestations sont échantillonnées à intervalle de 10 ans.
			Statistiques d'Exploitation du réseau TAC	5	TPG	12.21 E	
11	Hastus		Principes de conduite	10	TPG	7.01 E	
			Analyse des temps de parcours (ATP)	5	ECH	12.23 E	ATP à conserver pour les différents types de véhicules tous les 5 ans.
			Liste des journées	10	TPG	12.25 E	Conservation uniquement de la

							version de l'étape de décembre.
			Liste des voitures	10	ECH	12.26 E	Un exemplaire conservé tous les 5 ans.
			Horaires conducteurs, journée des roulements conducteurs	10	TPG	12.27 E	Conservation uniquement de la version de l'étape de décembre. Egalement dans intranet
			Horaires et rotations CTT, rotation SR, rotations RG	5	TPG	12.27.1 E	Conservation uniquement de la version de l'étape de décembre.
12	Intranet	Page conduite	Description des roulements conducteurs	10	TPG	7.02 E	
		Page conduite	Manuel du conducteur	5	TPG	1.13 E	
		Page conduite	Manuel d'exploitation	5	TPG	1.13.1 E	
		Doc de références	Guide de conduite des APOP	5	TPG	7.34 E	
			Horaires conducteurs, journée des roulements conducteurs	10	TPG	12.27 E	Conservation uniquement de la version de l'étape de décembre.
			Liste des arrêts	5	TPG	12.28 E	Année en cours sur intranet
			Code de parcours et destination	5	TPG	12.29 E	Année en cours sur intranet
			Liste des Hauts-le-pied	5	TPG	12.31 E	Conservation uniquement de la version de l'étape de décembre.
			Revue de presse	5	TPG	4.02 DG	

		Manuel de référence comptable	10	TPG	1.04 FG	
		Politique du personnel	10	TPG	7.05 RH	Aussi dans la GED
		Charte graphique	0	TPG	1.02.1 MVC	
		Rapport annuel de gestion	10	TPG	1.07.2 MVC	Aussi sur lecteur H/Com et site internet
		Info (selon le DA concerné)	5	TRI	8.15 MVC	*Ne conserver que les pièces de nature décisionnelle contentieuse ou organisationnelle.
13	Isilog	Manuel de qualité	10	TPG	1.29 DG	
		Rapports d'audit sur le fonctionnement financier et organisationnel	10	TPG	1.18 CA	
		Fiche d'amélioration, suivi des audits	10	TRI	1.15 DG	*Tri qualitatif des fiches selon le critère suivant : conservation des fiches qui découlent d'un audit et auxquelles le rapport d'audit est annexé.
		Rapports d'audit des autorités de surveillance	10	TPG	1.34 DG à 1.46 DG	Dans Isilog avec la fiche d'amélioration associée; le rapport est aussi dans la GED
		Rapport d'événement suite à un accident professionnel	10	ECH	7.09 DG	Conservation Rapports des personnes dont le dossier RH est échantillonné (cadre. Cadre supérieur, directeur et nom de famille commençant

							par lettre B). A la fin du DUL, les rapports sont intégrés aux dossiers du personnel.
			Carte stratégique	10	TPG	1.21 DG	Aussi dans la GED
15	MailChimp		Lettre d'information (newsletter)	5	TPG	10.10.1 DG	
15a	One Note		Direction F&G: dossiers de travail	10	TPG	12.01 FG	également dans le disque H
16	QlikSense		Statistiques générales et particulières des services, statistiques nombre d'incidents par type de véhicule	5	TPG	1.20 E, 1.21.1 E	Moyenne sur l'année à conserver
17	SAP		Plan comptable	10	TPG	6.04 FG	
			Assermentations	10	TPG	7.21 RH	Liste annuelle
			Dossiers du personnel permanent	10	ECH	7.27 RH	Versement aux archives TPG des dossiers des cadres supérieurs, des cadres, des directeurs et du personnel dont le nom débute à la naissance par la lettre B.
			Dossiers du personnel actif à durée déterminée	10	ECH	7.29 RH	Versement aux archives TPG des dossiers des cadres supérieurs, des cadres, des directeurs et du personnel dont le nom débute à la naissance par la lettre B.
			Dossiers des stagiaires	10	ECH	7.31 RH	Versement aux archives TPG des dossiers des cadres

							supérieurs, des cadres, des directeurs et du personnel dont le nom débute à la naissance par la lettre B.
			Dossiers des apprentis TPG	10	ECH	7.32 RH	Versement aux archives TPG des dossiers des cadres supérieurs, des cadres, des directeurs et du personnel dont le nom débute à la naissance par la lettre B.
			Statistique du personnel	5	TPG	7.41 RH	aussi dans disque H:\RH\SIRH
18	Site internet tpg		Communiqués de presse et points presse	10	TPG	4.09 DG	
			Rapport annuel de gestion	10	TPG	1.07.2 MVC	Aussi sur lecteur H/Com et intranet
19	Smart cockpit			10	TPG	1.14.1 DG	Comment se présentent les données? Tout archiver?
			Indicateurs de performance	5	TPG	1.15 CA	
20	Solu-Qiq		Organigrammes	10	TPG	1.01 DG	
			Cartographie des processus	10	TPG	1.22 DG	Conservation de tous les processus lors de chaque certification (tous les 3 ans)
			Procédures	10	TPG	1.24 DG	
			Fiche caractérisation processus	10	TPG	1.23 DG	
			Recueil de procédures comptables - Système de contrôle financier interne	10	TPG	6.02 DG	A vérifier

			Manuel SMI	10	TPG	1.29.1 DG	
			Plan commun d'audit interne	10	TPG	1.31 DG	GED et Soluquiq
21	Teams		Cellule de crise	10	TPG	1.27 DG	Transférer tout ce qui est dans teams dans eDocs (GED)?
			Réunion d'Exploitation	10	TPG	8.01 E	
			Cahier des charges de la ligne	10	TPG	12.19 E	Egalement dans Confluence?
22	tpgConnect		Dossier de recours de 2ème niveau (CIR)	10	TPG	8.14 DG	
23	Serveurs de fichier/disque réseau (P, Q, H)	Lecteur U/Interviewé 15	Discours, allocution, événement officiels, conférence	10	TPG	10.02 DG	
		Lecteur H/Com	Photographies d'entreprise	5	TPG	10.07 MVC	
		Lecteur H/Com	Publications des TPG (communication interne)	10	TPG	4.05 MVC	
		Lecteur H/Com	Rapport annuel de gestion	10	TPG	1.07.2 MVC	Aussi sur intranet et site internet
		Lecteur Q/ Division/Doc_Entr/NI	Notes d'information (NI)	5	TRI	8.07 E	*ne conserver que les pièces de nature décisionnelle contentieuse ou organisationnelle.
		Disque P/Administrateur CA	Délégation En Chardon	10	TPG	8.08 CA	
		Disque P/Administrateur CA	Société du Téléphérique du Salève	10	TPG	9.10 CA	

		Disque P/Administrateur CA	TPG - France	10	TPG	9.11 CA	
		Disque P/Administrateur CA	TP2A	10	TPG	9.12 CA	
		Lecteur H/F&G -Cge	Budgets annuels de fonctionnement et d'investissement	10	TPG	6.09 FG	
		Lecteur H/F&G -Cge	Budget de la présidence	10	TPG	6.01 CA	
			Synthèse analytique des dettes et des immobilisations	10	TPG	6.36 FG	
		Lecteur J	Publications produites par les TPG (communication externe)	5	TPG	4.03 MVC	
		Lecteur P	Horaires web, affiches horaires	5	TPG	4.04 MVC	
		Lecteur P/Présidence	Dossiers de travail de la présidence	10	TPG	12.01 CA	
		Lecteur P	Concept de sûreté	10	TPG	1.48 DG	
		Disque H/Direction F&G	Direction F&G: dossiers de travail	10	TPG	12.01 FG	Et également sur One Note
		Disque H/RH	Description de fonction - modèle de description de poste	5	TPG	7.13 RH	
		Disque H/RH	Description de poste	5	TPG	7.14 RH	
		Disque H/RH	Echelle des salaires et tableau des primes	30	TPG	7.23 RH	
		Disque H/RH	Statistiques formation	5	TPG	7.35 RH	H:\RH\Développement organisationnel\Unité développement et formation

					H:\RH\Développement organisationnel\Unité développement et formation\Séances délégués
Disque H/RH	Séance délégués formation	10	TPG	8.02 RH	
Disque H/RH	Statistique du personnel	5	TPG	7.41 RH	H:\RH\SIRH, aussi dans SAP
Disque H/RH	Séance délégués formation	10	TPG	8.02 RH	
Disque H/RH	Direction : dossiers de travail du directeur RH	10	TPG	12.01 RH	H:/RH/RH direction 2022
Disque H/Cge	Dossiers de subventions cantonales	10	TPG	6.10 FG	Se trouve dans le budget (GED) et dans le Disque H
Disque Q	Séance OPS (Opérations) hebdomadaire Service client opérationnel	10	TPG	8.01 MVC	Q/DIVISION/C/SCO
Disque H	Dossier Médecine du travail	30	TRI	7.47 RH	H:\RH\QVT - Qualité de Vie au Travail\Médecin du travail
Disque Q	Plans des véhicules	vie véhicule	TPG	12.20.1 T	Q/Division/T/DAO
Disque Q	Dossier bâtiment: autorisation de construire	Vie bâtiment	TPG	2.03 IB	Q/Division/DRI/I&B
Disque H	Direction : dossiers de travail du directeur T	10	TPG	12.01 T	Et dans la GED
Disque U/M. Mutter	Discours, communication	5	TPG	10.01 CA	
Disque H	Réalisation trolleybus, sous-station, aiguilles électriques : phase d'exécution et dossier conforme à exécution	Durée ligne	TPG	12.05 IB	
Disque H	Schémas installations de sécurité ferroviaire / Inst. 600V / Sous-stations	durée IS	TPG	12.27 IB	Q/Division/T/DAO

		Disque I	Analyses / expertises de sécurité pour les installations de sécurité ferroviaire (IS)	Durée IS	TPG	12.20 IB	I/Electrontechnique de traction
		Disque I	Homologation et versions logiciels installations de sécurité ferroviaire	10	TPG	12.26 IB	I/Electrontechnique de traction
		Disque H	Projet tram, trolleybus, sous-station, aiguilles électriques : phase d'étude et d'autorisation de construire (fédérale), Procédure d'approbation des plans (rapport technique) et décision d'approbation des plans	Durée ligne	TPG	12.03 IB	
		Disque H/HGM	Horaire-guide	2	ECH*	12.30 E	*Echantillonnage : 1 classeur complet pour chaque type de véhicules, tous les 5 ans.
24	GED edocs		Dossier audit des sous-traitants	10	TPG	12.22.1 E	
			Instructions de service (IDS)	5	TPG	1.02 E	Valable pour tous les DA
			Séances direction Exploitation et partenaires sociaux	5	TPG	1.22 E	
			Dossier d'appels d'offres pour les marchés publics et dossier de l'adjudicataire	10	ECH	3.01 E	Echantillonnage selon les 2 règles suivantes : 1) Conservation des dossiers des marchés publics supérieurs à 1 million de CHF et attributions de sous-traitance d'exploitation de lignes. 2) Echantillonnage qualitatif des dossiers

					de projet qui ne sont pas adjugés par le CA, selon les critères suivants : visibilité médiatique, consultation des rapports d'activités, etc.
	Catalogue des formations	5	TPG	7.25 E	
	Programme de formation	5	TPG	7.26 E	
	Plan de cours	5	TPG	7.27 E	
	Support de formation/cours	5	TPG	7.28 E	Aussi 360 learning?
	Programme/plan de formation en vue de l'obtention du CAP conducteur	10	TPG	7.31 E	
	PV des séances effectifs	10	TPG	7.44.1 E	
	Séance du pôle management E	10	TPG	8.02 E	
	Séance management conducteur Jonction / Bachet	10	TPG	8.03 E	
	Séance formations Tramways / Pneus	10	TPG	8.04 E	
	Séance mangement sûreté	10	TPG	8.05 E	
	Cellule LDT	10	TPG	8.05.01 E	
	Rapport d'activités au Conseil d'administration	10	TPG	8.16 DG	
	Notes de synthèse, présentations faites en Réunion de direction	10	TPG	8.18 DG	
	Séance avec la Direction générale des Transports (DGT)	10	TPG	9.01 E	PV
	Comité technique v.com (vitesse commerciale)	10	TPG	9.02 E	PV
	Comité de pilotage sur la vitesse commerciale	10	TPG	9.03 E	PV
	Correspondance	10	TPG	11.01 DG	

					*L'identification des projets à échantillonner se fait par le PMO (Project Management Officer), le chef de projet, le mandant et l'archiviste. Les critères sont les suivants: projet à plus d'1 million, visibilité médiatique, impact sur le fonctionnement de l'entreprise, etc.
	Projet métier / entreprise	10	ECH	12.02 E	
	Rapport sécurité	5	TPG	12.10.5 E	
	Calendrier de l'application de l'offre	5	TPG	12.24 E	
	Procédures internes BEX	5	TPG	12.40 E	
	Procédures métier SIT	10	TPG	1.10 SIT	
	Audits externes du SIT	10	TPG	1.16 SIT	
	Politique de sécurité de l'information	5	TPG	1.22 SIT	
	Référentiel des services et des traitements	5	TPG	4.03 SIT	
	Séance du domaine d'activité, plénière	10	TPG	8.01 SIT	
	Séance management	10	TPG	8.02 SIT	
	Objectifs annuels d'entreprise	10	TPG	1.05 DG	
	Revue de direction	10	TPG	1.07 DG	
	Cellule de pilotage des projets	10	TPG	1.17 DG	
	Directives santé et sécurité au travail	10	TPG	1.18 DG	
	Politique de gestion de risques	10	TPG	1.19 DG	
	Charte d'entreprise	10	TPG	1.28 DG	
	Manuel de référence - Management de projet	10	TPG	1.20 DG	Et aussi sur intranet

	Rapport d'audit interne	10	TPG	1.33 DG	
	Politique de gestion documentaire	10	TPG	4.14 DG	
	Règlement de gestion documentaire	10	TPG	4.15 DG	
	Dispositions réglementaires de transport (DRT-TPG)	10	TPG	5.02 DG	
	Concessions fédérales d'infrastructure	10	TPG	5.01 IB	GED + Dir'AJ?
	Concessions fédérales d'exploitation de l'infrastructure (autorisation d'exploiter l'infrastructure)	10	TPG	5.02 IB	GED + Dir'AJ?
	Concession pour des transports réguliers de voyageurs effectués à titre professionnel au moyen de tramways	10	TPG	12.33 MVC	Aussi dans Dir'AJ
	Autorisation cantonale pour des transports réguliers de voyageurs effectués à titre professionnel au moyen de bus	10	TPG	12.34 MVC	Aussi dans Dir'AJ
	Concession de lignes pour les autobus	10	TPG	12.36 MVC	Aussi dans Dir'AJ
	Concession pour les trolleybus	10	TPG	12.37 MVC	Aussi dans Dir'AJ
	Homologation des séries de véhicules	10	TPG	12.77 T	
	Autorisation d'exploiter un véhicule	10	TPG	12.78 T	
	Actes notariés	10	TPG	5.20 DG	
	Réunion de direction	10	TPG	8.01 DG	
	Séance OETHand	10	TPG	8.06 DG	Aussi dans Dir'AJ
	Statistiques : suivi des AIMP	5	TPG	8.08 DG	Uniquement les récapitulatifs annuels
	Séance du Comité interne de reconsidération (CIR)	10	TPG	8.11 DG	
	Règlement interne du CIR	10	TPG	8.12 DG	
	Rapport d'activité du CIR	10	TPG	8.13 DG	
	Plan marketing	5	TPG	1.03 MVC	

	Rapport annuel de réalisation de l'offre	10	TPG	1.07 MVC	
	Dossiers de subventions fédérales d'exploitation pour le trafic régional de voyageurs	10	TPG	6.02 MVC	
	Correspondances et séances de coordination pour des projets de développement du réseau	10	TPG	9.02 MVC	
	Plans de communication	10	TPG	10.04 MVC	
	Etude d'image de marketing / sondage / enquête	10	TPG	11.03 MVC	GED?
	Direction MVC: dossiers de travail	10	TPG	12.01 MVC	
	Rapport d'activité du service relation clientèle	10	TPG	12.16 MVC	Ne conserver que la version annuel
	Plan d'action des transports collectifs (PATC)	10	TPG	12.29 MVC	
	Rapport technique	10	TPG	12.30 MVC	
	Rapport d'étape	10	TPG	12.31 MVC	
	Concession pour des transports réguliers de voyageurs effectués à titre professionnel au moyen de tramways	10	TPG	12.33 MVC	Aussi Dir'AJ?
	Autorisation cantonale pour des transports réguliers de voyageurs effectués à titre professionnel au moyen de tramways	10	TPG	12.34 MVC	Aussi Dir'AJ?
	Rapport technique externe	10	TOG	12.39 MVC	
	Rapport annuel auprès de l'OFT	10	TPG	1.07.1 MVC	
	Séances de suivi du plan d'action des transports collectifs (PATC) et autres projets de développement du réseau	10	TPG	9.01 MVC	
	Autorisation de circuler en France	10	TPG	12.38 MVC	
	Répartition des tâches entre le président du CA et le DG.	10	TPG	1.03 CA	

	Répartition des tâches entre Risque & Management et Audit interne	10	TPG	1.04 CA	
	Règlement de gouvernance d'entreprise des TPG	10	TPG	1.05 CA	
	Règlement de la commission d'audit	10	TPG	1.06 CA	
	Règlement de la rémunération du collège de direction	10	TPG	1.07 CA	
	Séance du Conseil d'administration (CA)	10	TPG	8.01 CA	
	Séance du bureau du CA (BCA)	10	TPG	8.01 CA	
	Commission Audit & Finance (CAuFin)	10	TPG	8.03 CA	
	Commission Réseau et Infrastructures	10	TPG	8.04 CA	
	Commission ressources humaines et rémunérations	10	TPG	8.05 CA	
	Commission stratégique	10	TPG	8.06 CA	
	Séance avec le département de tutelle	10	TPG	9.06 CA	
	Programme financier relatif à l'entretien, au renouvellement et à la construction des infrastructures Tramway et Trolleybus	10	TPG	6.03 IB	
	Séance Infrastructures & bâtiments	10	TPG	8.01 IB	
	Séance Electriciens réseau 600 Volts	10	TPG	8.02 IB	
	Séance Voie et 600 Volts	10	TPG	8.03 IB	
	Séance Controlling infrastructures	10	TPG	8.05 IB	
	Séance bâtiment	10	TPG	8.08 IB	
	Comité de pilotage (Copil) et comité de direction (Codir) En Chardon	10	TPG	8.09 IB	
	Comité consultatif pour l'accessibilité des personnes à mobilité réduite	10	TPG	8.10 IB	
	Correspondance avec l'OFT	10	TPG	11.02 IB	
	Direction IB: dossiers de travail	10	TPG	12.01 IB	

	Règlements de travail sur les installations électriques ferroviaires (TB et TW)	10	TPG	12.19 IB	
	Support de formation cours (Exploitation / Technique) "600V"	10	TPG	12.25 IB	Vérifier si pas dans 360 learning
	Rapport de financement pour l'Etat (DF et DI)	10	TPG	6.06 FG	
	Rapport de financement pour le CA	10	TPG	6.06.1 FG	
	Statistiques des ventes de titres de transport	10	TPG	6.37 FG	Statistique annuelle à conserver
	Séance équipe de direction FG	10	TPG	8.01 FG	
	Statistiques : suivi des AIMP	5	TPG	8.08 FG	Statistique annuelle à conserver
	Séance syndicat SEV et Technique	5	TPG	1.06 T	
	Séance syndicat SEV et Cadres du Technique	5	TPG	1.07 T	
	Séance directeur et cadres supérieurs	10	TPG	8.01 T	
	Séance RH sur les collaborateurs techniques	10	TPG	8.02 T	
	Séance Electriciens réseau 600 Volts	10	TPG	8.07 T	
	Séance bâtiment	10	TPG	8.10 T	
	Correspondance avec l'OFT concernant les véhicules	10	TPG	9.01 T	
	Séance avec les constructeurs pour le suivi des garanties	Vie véhicule	TPG	9.02 T	
	Rapport de contrôle des mises à terre CMB (Centre de maintenance Bachet)	10	TPG	12.36 T	Sort final à confirmer
	Rapport de contrôle des mises à terre CMC (Centre de maintenance En Chardon)	10	TPG	12.37 T	Sort final à confirmer
	Rapport de contrôle des mises à terre des armoires d'alimentation	10	TPG	12.55.4 T	
	Etats financiers consolidés	10	TPG	6.25 FG	GED pour la version définitive (et H/Partage compta

					pour la version de travail)
	Comptes de résultat annuels	10	TPG	6.28 FG	GED pour la version définitive (et H/Partage compta pour la version de travail)
	Bilans annuels	10	TPG	6.29 FG	Conservé avec les Etats financiers consolidés 6.25.
	Plan commun d'audit interne	10	TPG	1.31 DG	GED et Soluquiq
	Procédure d'audit interne	5	TPG	1.13 CA	#1338
	Tableaux des indicateurs	5	TPG	1.15 CA	
	Charte de l'audit interne	10	TPG	1.17 CA	#1336
	Réunion de l'UTP (Union des transports publics) et Alliance SwissPass	10	TPG	1.57 DG	
	Règlement interne sur le paiement des jetons de présence aux administrateur des TPG	10	TPG	8.07 CA	GED + Lecteur P/Administrateurs
	Correspondance avec le Conseil d'Etat	10	TPG	9.01 CA	
	Arrêtés du Conseil d'Etat	10	TPG	9.02 CA	
	Correspondance avec le Grand Conseil	10	TPG	9.04 CA	
	Correspondance avec le département de tutelle	10	TPG	9.05 CA	
	Correspondance officielle	10	TPG	11.01 CA	
	Portefeuille de projet d'entreprise	5	TPG	1.10.01 DG	GED et confluence
	Rapport de gestion de risques	10	TPG	1.20.1 DG	
	Certification	10	TPG	1.14 DG	
	Politique environnementale	10	TPG	1.29.2 DG	
	Cellule de sécurité, sûreté	10	TPG	1.47 DG	

	Système de gestion de la sûreté, la sécurité et la santé au travail	10	TPG	1.49 DG	
	Politique de sûreté pour les infrastructures bâtiment	10	TPG	1.50 DG	
	Manuel sécurité	10	TPG	1.53 DG	
	Manuel de gestion de crise	10	TPG	1.54 DG	
	Séance et statut ESITPG (équipe de sécurité incendie TPG)	10	TPG	8.03 DG	
	Séance Secouriste	10	TPG	8.04 DG	
	Séance répondants risques	2	TPG	8.05 DG	
	Séance répondant qualité	2	TPG	8.05.1 DG	
	Rapports d'audit des autorités de surveillance	10	TPG	1.34 DG à 1.46 DG	Dans Isilog avec la fiche d'amélioration associée; le rapport est aussi dans la GED
	Arrêté départemental	10	TPG	9.03 CA	
	Norme ISO 21000	10	TPG	1.40 DG	
	Séance du CAD (Commission des achats durables)	10	TPG	8.07.1 DG	
	Plan financier pluriannuel (PfP)	10	TPG	1.08 DG	Aussi dans la Dir'AJ
	Plan financier quadriennal (pfq)	10	TPG	1.03.1 FG	
	Séance du CAD (Commission des achats durables)	10	TPG	8.07 FG	
	Réunions avec les syndicats, accords obtenus et PV de séances	10	TPG	1.06 RH	
	Statut du personnel	10	TPG	7.01 RH	
	Règlement d'application du statut du personnel	10	TPG	7.02 RH	
	Règlements particuliers	10	TPG	7.03 RH	

	Dossier de négociation des statuts	10	TPG	7.04 RH	
	Politique du personnel	10	TPG	7.05 RH	Aussi dans intranet
	Réunion interne des RH (RIRH)	10	TPG	8.01 RH	
	Renouvellement des membres des commissions statutaires	10	TPG	8.03.1 RH	
	Election des représentants du personnel au CA (TPG, FPTPG, etc.): PV de dépouillement signé	50j	TPG	8.11 RH	
	Carte stratégique	10	TPG	1.21 DG	Aussi isilog
	Règlement "piquet électriciens"	10	TPG	12.34 T	
	Règlements de travail sur les installations électriques ferroviaires (TB et TW)	10	TPG	12.47 T	
	Séance de la commission santé et sécurité au travail	5	TPG	8.05 RH	
	Rapport de contrôle des mises à terre des armoires de pontage	10	TPG	12.46.4 T	Sort final à confirmer
	Dossiers de travail du directeur SIT	10	TPG	12.01 SIT	
	Séance de la commission des véhicules	5	TPG	8.07 RH	
	Séance avec l'Office cantonal des Transports (OCT)	10	TPG	9.01 IB	
	Lien TPG & communes	10	TPG	9.03 IB	
	Correspondance et séances de coordination pour des projets de développement du réseau	10	TPG	9.04 IB	
	Autorisation d'accès au réseau	10	TPG	12.07 IB	
	Agrément de sécurité (AgSec)	10	TPG	12.10 IB	
	Certificat de sécurité (CerSec)	10	TPG	12.11 IB	
	Directives du Département des finances	10	TPG	6.02 FG	

		Déclarations signées rétrocession de la taxe carburant (RTC) et redevance poids lourds liée aux prestations (RPLP)	15	TPG	6.05 T	Uniquement la demande annuelle
		Dossiers de travail du directeur E	10	TPG	12.01 E	
		Instances faïtières des transports publics transfrontaliers (correspondance, PV)	10	TPG	1.23 E	
		Déclaration d'applicabilité	10	TPG	1.25 SIT	
		Dossier de travail du directeur général	10	TPG	12.01 DG	
		Dossier historiques et d'organisation des commissions statutaires	10	TPG	8.03 RH	GED?
		Séance de la commission des horaires	5	TPG	8.06 RH	GED?
		Séance de la commission circulation	5	TPG	8.08 RH	GED?
		Séance de la commission conciliation	5	TPG	8.09 RH	GED?
		Commission électorale	5	TPG	8.10 RH	GED?
	Stockage inconnu pour l'instant					
		Plan d'adaptation à l'évolution des nouvelles technologies	5	TPG	1.21 SIT	
		Aide mémoire relatif à la procédure disciplinaire	5	TPG	7.45 E	
		Dossier de réformes structurelles	10	TPG	1.13 DG	
		Campagne promotionnelle	10	TPG	12.27 MVC	
		Réalisations de campagne	10	TPG	12.28 MVC	

Annexe 4 : Guide d'entretien

Guide d'entretien

Cycle de vie des documents et dossiers dans les systèmes métier des tpg

Informations techniques

Nom de l'enquêteur : Adrian Rey Rodriguez

Type d'entretien : Semi-directif

Lieu de l'entretien : Siège des tpg à Bachet (Route de La-Chapelle 1, 1212 Grand-Lancy)

Date : jj.mm.aaaa

Personne ou groupe interrogé :

Logiciel(s) concerné(s) :

Durée estimée : 40 minutes

Contexte

Les archives des tpg ont pour mission de traiter non-seulement les archives papier mais aussi les archives électroniques. Les logiciels métier en usage aux tpg sont nombreux mais peu d'informations sur leur fonctionnalités et contenus sont connues du service des archives. Ce guide d'entretien a été créé pour mieux cerner les pratiques documentaires des collaborateurs et collaboratrices et le processus de gestion des documents dans ces logiciels.

Les archives veulent savoir :

- ➔ Comment le logiciel métier fonctionne.
- ➔ Si et comment le logiciel gère le cycle de vie des documents.
- ➔ Comment extraire les documents/dossiers à valeur historiques conservés dans les logiciels pour les verser dans la plateforme de pérennisation et d'archivage à long terme.

Questions

1. Profil de l'intervenant et description du logiciel métier (Réfèrent métier et SIT)

Présentation du référent métier et/ou SIT :

- 1.1. Pourriez-vous vous présenter (nom, prénom) et ainsi que décrire votre activité au sein des tpg ?
- 1.2. Pourriez-vous préciser quelles sont les tâches que vous effectuez en rapport avec le logiciel métier ?
- 1.3. Quelle est votre fréquence d'utilisation du logiciel métier ? (Tous les jours, quelques jours par semaine, un fois par semaine, quelques fois par mois, 1 fois par mois, rarement)

Présentation du logiciel métier :

- 1.4. De quel type de logiciel métier s'agit-il ? (GED, BDD, outil de communication, logiciel collaboratif...) ?
- 1.5. À quoi sert ce logiciel métier ? Quelle est sa fonction ou son utilité ?
- 1.6. Quelle est l'ancienneté du logiciel ? depuis combien de temps est-il en service ? Y a-t-il des mises à jour ?

2. Cycle de vie des documents/dossiers dans le logiciel (réfèrent métier)

- 2.1. Quel est le parcours du document/dossier type depuis sa création à sa version finale dans le logiciel ? Quelles sont les phases de la vie du document ?
- 2.2. Y a-t-il la possibilité de classer un document ou dossier par son statut d'utilisation ?
- 2.3. Le logiciel permet-il de catégoriser les documents selon leur cycle de vie/statut d'utilisation ? (Actif - utilisation courante, semi-actif – utilisation occasionnelle, inactif - version définitive ou qui n'est plus utile pour le maintien des activités actuelles)
- 2.4. Est-il possible d'assigner des durées de conservation à un document ? Par exemple, le logiciel peut-il indiquer que les 10 ans liés à la conservation d'une facture sont écoulés ? (DUL)
- 2.5. Le logiciel permet-il d'éliminer et/ou de transférer automatiquement les documents ou dossiers vers un environnement extérieur ? Y a-t-il une fonctionnalité qui permette, par exemple, de transférer des séries de documents directement vers un autre outil de gestion ?

- 2.6. Que faites-vous lorsqu'un document ou un dossier atteint sa version finale ou qu'il est complet et qu'il n'est plus utilisé ?
- 2.7. Avez-vous déjà éprouvé des difficultés à trouver ce que vous cherchez parmi la masse d'informations ?
- 2.8. Renseignez-vous des métadonnées (informations) à propos du document ou dossier ?
Si oui, quelles métadonnées renseignez-vous ? S'agit-il de métadonnées décrivant le contenu ou de métadonnées avec des informations à propos de la préservation ? À votre avis ce renseignement est-il suffisant où y a-t-il des lacunes ?

Si non, pourquoi ? Y a-t-il des métadonnées qui pourraient être pertinentes ?
- Le logiciel assigne-t-il automatiquement des informations à propos d'un document ? (Par exemple : peut-il reprendre automatiquement le titre d'un document et l'ajouter au champs « titre » des métadonnées) ?*
- 2.9. Y a-t-il des données personnelles ou sensibles ? Est-ce que le logiciel peut gérer le niveau de sensibilité des documents ? Les indications à propos du niveau de sensibilité indiquées dans le référentiel sont-elles à jour ?
Le référentiel classifie [nom logiciel] comme S ? (données sensibles/confidentielles), est-ce que cela concerne aussi les archives historiques listées ?

3. Conservation dans le logiciel métier (référent SIT)

- 3.1. Le logiciel garantit-il la valeur probante des dossiers et documents ?
Peut-on vérifier si la version du document est l'originale ou s'il a été modifié ? Est-il probant au niveau interne des tpg voir au niveau de la justice ? Permet-il de procéder à un horodatage certifié et à de la signature électronique avec un niveau élevé ?
- 3.2. Quels sont les formats gérés par le logiciel (PDF, XLS, DOCX...) ?
Quel est le plus courant ?
- 3.3. Avez-vous déjà été dans l'incapacité d'ouvrir un fichier ?
Si oui, le fichier était corrompu ? le format n'était plus supporté ? l'affichage était déformé ? le fichier n'existait plus ?
- 3.4. Quelle est la volumétrie du logiciel ?
Avez-vous constaté des problèmes liés à la volumétrie des contenus ? (Surcharge d'informations ? Espaces de stockage saturés ?)

Quel est le volume des documents/dossiers à valeur historique ?
- 3.5. À l'aide de la liste des séries de documents ou dossiers historiques que nous vous avons transmis pour ce logiciel, pouvez-vous nous expliquer comment chacun d'eux est conservé ?
- 3.6. Parmi les documents historiques listés pour ce logiciel métier, y a-t-il des contenus aux formats atypiques ?

4. Extraction des documents/dossiers (réfèrent SIT)

Le service d'archives souhaiterait à terme, extraire les documents ou dossiers à valeur historique se trouvant dans le logiciel métier vers une plateforme de pérennisation et d'archivage à long terme. Le service d'archives souhaiterait donc pouvoir extraire un paquet d'information contenant le contenu d'information (données et/ou documents) et l'information de description et de pérennisation (métadonnées).

- 4.1. Quel serait le processus/moyen d'extraction le plus simple ? Est-ce qu'une API conviendrait ?
- 4.2. Quels seraient les moyens concrets pour définir une interface de programmation (API) ? Quelles sont les démarches et prérequis ?
- 4.3. Quels seraient les difficultés et obstacles prévisibles à l'extraction ? Y a-t-il des types de contenu qui pourraient subir une altération de leur qualité visuelle lors d'une migration de format ou d'une extraction ?

Remerciements et demande d'une brève démonstration.

Annexe 5 : Tableau synthétique des entretiens

Partie 1

	360 Learning	Allegro	Confluence	Dir'AJ	GestE	Hastus
1. Profil et description						
1.1. Réf. Métier : Présentation	Interviewé 1 (DG), Gestionnaire des risques d'entreprise et Responsable du Management d'entreprise. Fait partie d'une équipe qui aborde les thématiques SQSE (santé, sécurité, sûreté, qualité), gestion du portefeuille de projet, la gestion documentaire et l'environnement.	Interviewé 3 (RH), ancien Responsable SIRH, GT et Sal. Actuellement intérim au pôle Administration et basé dans la cellule analytique. (Le référentiel services et traitement ne serait pas à jour).	Interviewé 1 (DG), Gestionnaire des risques d'entreprise et Responsable du Management d'entreprise. Fait partie d'une équipe qui aborde les thématiques SQSE (santé, sécurité, sûreté, qualité), gestion du portefeuille de projet, la gestion documentaire et l'environnement.	Interviewé 6 est Responsable achats sous F&G. Il y a deux entités au niveau Responsable achats : les acheteurs et le back-office achat dont l'interviewé 7 fait partie. L'interviewé 7 est Gestionnaire achats sous F&G. Le back-office achat fait toute l'administration et gère toute la partie DirAJ.	Interviewé 9 (SJ), Gestionnaire des sinistres. S'occupe de tous les accidents qu'il y a sur le réseau et du portefeuille d'assurance des tpg.	Interviewé 11, E, Responsable Pôle Conception de l'offre. A comme objectif d'avoir un horaire pour un arrêt avec un conducteur et un véhicule.
1.2. Réf. Métier : Tâches en lien avec l'app. métier	N'utilise 360 Learning qu'en tant que client utilisateur pour suivre les formations proposées.	Devait générer des APOP, faire les APOP et les valider.	Pas de tâches en rapport avec Confluence.	Pour F&G, DirAJ permet surtout la gestion de contrats d'achats, contrats de prestations et de partenariats avec des communes et événements. L'interviewé 7 s'occupe de suivre le cycle de vie des contrats d'achats.	Fait l'enregistrement des dossiers sur GestE. GestE met en relation avec les estimateurs pour avoir les devis de réparation responsabilité civile. Tous les documents nécessaires sont enregistrés sur GestE.	Créé les différents horaires pour les clients, les véhicules et les conducteurs à l'aide de Hastus.

1.3. réf. Métier : Fréquence d'utilisation	Rarement/Voir réponse 1.2.	Rarement, quelques fois par année lors de la période des APOP.	N'utilise pas Confluence. C'est principalement l'IT et le Responsable du PMO qui l'utilisent.	Utilisation quotidienne pour l'interviewé 7 et utilisation en mode consultation quelques fois par semaine pour l'interviewé 6.	Utilisation quotidienne, il s'agit de l'outil principal pour ce travail.	Tous les jours.
1.1. réf. SIT : Présentation	Interviewé 2, SIT, Ingénieur d'application IT. A la responsabilité technique de l'application 360 Learning.	Interviewé 4, SIT, Cheffe de projet IT. Administratrice d'Allegro avec une personne en RH. Ne s'occupe pas du développement du logiciel. Allegro a été acheté à son fournisseur (Adequasys). Il s'agit d'une vieille version qui est maintenue en attente d'une optimisation de SAP au niveau des RH.	Interviewé 5, SIT, Chef de projet IT. Gère la partie des solutions, le système de vente et toutes les applications qui en découlent : WebShop etc.	Interviewé 8 (SIT), Chef de projet IT. Gère le cycle de vie logiciel, répond aux besoins de solutions et teste les nouvelles solutions.	Interviewé 10 (SIT) Cheffe de projet IT, est responsable applicative et gère le cycle de vie, les besoins métiers et les évolutions. S'occupe également de toutes les solutions autour de la téléphonie entre les systèmes de sécurité/alarme.	Interviewé 12, SIT, Cheffe de projet IT. S'occupe de l'aspect fonctionnel, de la relation avec le fournisseur et du cycle de vie du logiciel. "Le logiciel doit fonctionner en permanence sinon aucun bus ou tram ne peut rouler."
1.2. réf. SIT : Tâches en lien avec app.	Référent pour tous les problèmes ou évolutions, se charge du contact avec le fournisseur.	Maintien du logiciel Allegro.	Met à jour les outils quand il y a des vulnérabilités applicatives, réponds aux questions etc. Met en place différentes standards, par exemple, la stratégie des droits des accès.	L'interviewé 8 s'occupe de faire les montées de version de Diraj lorsqu'il y a des besoins exprimés par F&G et gère le cycle de vie du logiciel.	Gère les soucis rapportés sur GestE (sauf les problèmes liés aux serveurs).	Il faut gérer des questions métier, des analyses, les anomalies, les demandes d'amélioration, des demandes de corrections.
1.3. réf. SIT : Fréquence	Une ou quelques fois par semaine.	Rarement.	Quelques fois par semaine. (représente 20% du travail par semaine)	Tous les jours.	Tous les jours. GestE est surtout utilisé par la RCT qui est le point d'entrée au niveau des événements.	Tous les jours.

1.4. Type d'app.	360 Learning est un LMS (Learning Management System). Les tpg l'utilisent plutôt comme une plateforme de formation digitale. Il s'agit d'un client léger.	Allegro est un logiciel de gestion RH. C'est un client léger qui s'affiche sur le navigateur web.	Interviewé 1 : Confluence est comme un Wiki Tool qui permet de pouvoir sectoriser de l'information et donner des accès spécifiques. Interviewé 5 : Confluence est comme Wikipédia, c'est un outil pour faire de la documentation collaborative : c'est un format Wiki facile à lire et qui est standardisé.	DiraJ est un logiciel juridique. C'est tant un logiciel de gestion juridique qu'une base de données.	GestE veut dire : Gestion des événements réseau. C'est une solution centralisée des événements qui assurent notre réseau.	Hastus est un logiciel de gestion et de planification du transport public. « C'est le SAP du transport public ».
1.5. Fonction app.	360 Learning permet de digitaliser certaines formations.	Allegro sert à remplir les évaluations régulières des collaborateurs tpg. Il permet de générer des APOP, de les remplir et les valider par les différents responsables. Cela permet d'avoir un feedback des collaborateurs.	Confluence sert à partager de l'information de manière collaborative au sein de l'entreprise sous le modèle de Wikipédia. Plusieurs personnes peuvent enrichir une documentation à propos de, par exemple, l'utilisation d'une application et apporter des informations complémentaires lors de changements dus à une mise à jour.	DiraJ permet de gérer les contrats et fonctionne également comme une base de données des contrats.	GestE permet d'administrer toutes les tâches liées aux sinistres. Il permet de centraliser toutes les interventions des différents acteurs concernant un événement (par exemple : un accident de circulation).	Hastus permet de concevoir la production de l'offre tpg, c'est-à-dire la combinaison de l'heure, de l'arrêt, du véhicule et du conducteur. Hastus gère aussi tous les plannings des conducteurs, les repos, les vacances, etc.

1.6. Ancienneté app.	360 Learning existe depuis fin 2018, il y a des mises à jour régulières.	Allegro existe depuis 2010 aux tpg, les mises à jour sont peu fréquentes. Avec l'évolution prochaine de SAP S/4, des questions concernant Allegro vont être relancées. (Commentaire de l'intervisé 4 : Dans le futur peut-être que SAP ou un autre outil RH sera utilisé à place d'Allegro).	Existerait depuis plus de 6 ans pour l'Exploitation et depuis quelques années pour SIT.	DiraJ (version actuelle) est introduit en 2014, il y a des mises à jour régulières. Avant 2014 DiraJ existait sous une autre version, il s'agissait d'une application « dure » alors qu'aujourd'hui il s'agit d'une application web (client léger).	GestE existe depuis juillet 2014 aux tpg. Avant, la centrale de régulation avait son propre outil : Journal Réseau. Le Service juridique avait Fox Pro. Le Technique ne travaillait qu'avec SAP. Les communications se faisaient par e-mail.	Existe depuis 1998. La grande dernière mise à jour date de fin mars 2022.
2. Cycle de vie						
2.1. Parcours/phases de vie	Un formateur peut créer un support de formation directement sur 360 Learning mais il peut aussi téléverser des documents (vidéos tutoriels, PowerPoints etc.). Le formateur liste les collaborateurs potentiels pour la formation et ceux-ci reçoivent un mail leur recommandant la formation. Il y a des petits quizz et les collaborateurs peuvent remplir des feedbacks sur la formation. Il est possible de voir l'aperçu de la moyenne des résultats globaux pour une formation.	Les APOP sont créées sur Allegro, elles sont alors « ouvertes » par le responsable qui peut renseigner certains éléments. Il y a tout ce qui est lié au courant (l'évaluation en cours), les objectifs reviennent de la dernière APOP afin de les évaluer et les objectifs futurs sont demandés. L'APOP est validée par le collaborateur, le responsable, le supérieur et les RH (Pôle Conseil). Les RH reçoivent probablement une notification ou pas que le document est en attente de validation de leur part. Il n'y a pas de récupération automatique. Pour valider,	Interviewé 1 : Il n'y a pas de documents créés sur Confluence. Tous les documents qu'il y a sur Confluence sont téléversés (uploadés). Il y a des pages textuels faisant partie de l'affichage de Confluence et il est possible de retrouver certaines images cliquables avec des liens hypertextes (Similaire aux organigrammes cliquables de l'intranet).	Les contrats originaux sont reçus en version papier. Le contrat papier est numérisé en version PDF simple. DiraJ est composé par dossiers, le contrat est rattaché si nécessaire à des pièces-jointes. Des métadonnées sont entrées manuellement. Lorsque la validité du contrat arrive à terme, il est clos et stocké dans DiraJ. La version papier est alors archivée physiquement.	L'enregistrement d'un cas de sinistre se fait directement sur GestE, dans le service tout le monde a accès à la modification. Selon la particularité ou gravité d'un cas, sa clôture peut prendre entre une semaine (pour accident léger) à plusieurs années (pour un accident impliquant des blessés etc.).	Plusieurs des documents listés sont aussi sur intranet. Les documents sont créés directement sur Hastus. Ils sont extraits d'Hastus au format PDF. Une fois qu'ils sont validés, ils sont mis à disposition des conducteurs sur intranet.

		les RH passent en revue l'APOP afin de détecter un possible problématique.				
2.2. Assignment statut manuel	Non/Ne s'applique pas.	Voir réponse 2.3.	Non/Ne s'applique pas.	Un contrat est soit ouvert ("actif" sur Diraj) ou soit clos. (Constatation d'un onglet déroulant pour indiquer le statut du document lors de la démonstration).	Les évènements peuvent être classés sous divers statuts : déclaré, transmis etc. GestE ne peut qu'indiquer que si un dossier est ouvert ou s'il est clôturé.	Un document est d'abord "en cours de travail" puis il est mis "en production" lorsqu'il est validé.
2.3. Assignment statut automatique	Non/Ne s'applique pas.	Allegro indiquerait aux RH si une APOP est en attente de validation.	Non/Ne s'applique pas.	Non.	Non.	Non.
2.4. Assignment durées de conservation	Non/Ne s'applique pas.	Non, Il n'y a pas de notion de type calendrier de conservation.	Non/Ne s'applique pas.	Il y a la possibilité de lancer des rappels/alertes, par exemple, à 4 mois de l'échéance de validité d'un contrat.	Le développement d'un système pour assigner des durées de conservation n'existe pas mais serait souhaitable. Un système d'agenda (non lié à GestE) permet actuellement de créer des rappels pour, par exemple, relancer une personne dans un recours après deux mois.	Non.

2.5. Élimination/tr ansfert automatique/ interconnexio ns	Non/Pas d'informations.	Il y aurait une connexion SAP-Allegro pour transférer les données de la personne.	Non, mais certains documents de Confluence sont liés à la GED eDocs.	Pas d'éliminations. Mais il y aurait des connexions avec le SAP.	Il est possible de transférer à l'intérieur de GestE vers le système GestE des estimateurs par exemple. Il ne serait cependant, pas de fonctionnalité permettant de transférer de GestE vers un autre logiciel métier.	Non.
2.6. Pratiques version finale	Ne sait pas si les anciennes versions d'une formation sont éliminées et remplacées par la nouvelle, modifiées ou bien conservées. Ne sais pas comment les données sont stockées.	Les APOP restent sur Allegro.	Interviewé 1 : Tout est gardé. Les documents liés à eDocs vivent leur cycle de vie en fonction de eDocs. Toutes les versions d'un document sont gardées. Interviewé 5 : Cela dépend du souhait du propriétaire de la page. Mais il est possible de réécrire par-dessus et donc d'effacer l'ancien contenu.	Les dossiers de contrats sont clos et stockés dans Diraj, la version papier est également conservée.	Quand un dossier ou document atteint sa version finale il est stocké dans GestE indéfiniment.	En général, les documents s'empilent puis à un moment, la Communication est tenue d'enlever tous les anciens qui sont alors éliminés de Hastus mais conservés sur un répertoire interne indéfiniment.
2.7. Difficultés de recherche	Non, l'outil de recherche est simple.	Non, il suffit de taper le nom ou la matricule de l'employé.	Non.	Pas de difficultés de recherche. Il est aussi possible de faire une requête pour obtenir, par exemple, la liste de tous les contrats avec les avenants qui ont une fin de validité pour l'année X. Diraj génère alors un tableau Excel avec la liste demandée.	Il arrive d'avoir des difficultés à retrouver un document. Il s'agit surtout d'un "défaut humain" puisque si les métadonnées ne sont pas correctement entrées par le collaborateur il s'avère difficile de retrouver le document par la suite.	Non.

2.8. Renseignements des métadonnées	Ne sait pas/À priori non.	Il n'y a pas d'informations quant aux métadonnées qui accompagneraient les documents extraits au format RTF.	Il n'y aurait pas de métadonnées renseignées mais les métadonnées d'un document téléversé seraient conservées.	Les métadonnées sont assignées manuellement. Il n'y aurait pas de lacunes dans les métadonnées, leur remplissage est assez protocolé et nécessaire au quotidien.	Les collaborateurs qui travaillent sur GestE renseignent eux-mêmes les métadonnées manuellement. Plusieurs champs de métadonnées sont nécessaires dépendant de la situation : noms et prénoms d'une personne lésée, numéro de sinistre, adresses, etc. Les seules métadonnées automatiques sont l'horodatage et le nom de la personne qui a introduit le dossier.	Non et il n'y a pas de métadonnées entrées manuellement dans les PDF générés.
2.9. Gestion des données sensibles	Le référentiel classifie 360 Learning comme S2 (données internes : informations relatives au fonctionnement des services)	Le référentiel classifie Allegro comme S0 (données sensibles/confidentielles).	Le référentiel classifie Confluence comme S2 (données internes : informations relatives au fonctionnement des services).	Le référentiel classifie DiraJ comme S0 (données sensibles/confidentielles).	Le référentiel services et traitements classifié GestE comme S0 (données sensibles ou confidentielles). Il existe sur GestE un onglet "Confidentiel". Il peut y avoir des documents S0 parmi ceux listés comme archives historiques. (Exemples de documents S0 : rapports de police, certificats médicaux de personnes lésées, etc.).	Le référentiel classifie Hastus comme S1 en (données personnelles non sensibles), cela concerne aussi les documents listés à valeur historique.

3. Conservation						
3.1. Garantie de valeur probante	Non.	Non/Ne s'applique pas. Mais seul les APOP conservées sur Allegro font foi.	Non. Pour chaque personne qui va éditer sur Confluence, l'avantage est qu'on a l'historique, on a le nom de la personne qui l'a fait et l'heure de la modification. Il y a aussi les versions, tout est tracé et il y a des notifications lorsque quelque chose est modifié.	Les contrats sont des copies PDF de leur original papier avec donc la signature papier numérisée. Il n'y a pas de contrats numériques et donc pas de signature électronique. Il n'y a pas d'horodatage certifié.	Il s'agit surtout d'emails reçus ou de correspondances scannées au format PDF. Pour le reste des documents de travail il est possible pour les collaborateurs de modifier un document.	Non/Ne s'applique pas.
3.2. Formats	On peut retrouver des MP4, PDF, PowerPoint ou autres documents de la suite Microsoft Office ainsi que quelques images (probablement Jpeg). Il y a aussi pas mal de code Embedded (code incorporé : programme ou un objet utilisé dans un autre programme.)	Les APOP ne sont pas dans un format particulier sur Allegro (elles font partie de l'affichage). L'APOP peut être exportée dans un format qui est modifiable : format .RTF qui peut être ouvert sur Word.	Il est possible de rattacher des pièces-jointes aux pages de Confluence : PDF, Word, formats de la Suite Microsoft Office, Visio. Les textes/pages sur Confluence sont au format Wiki.	Le PDF est le format le plus courant. Il peut y avoir des Word lorsque le contrat compte des pièces-jointes. Il y a aussi des Emails au format MSG.	Surtout du PDF, il peut y avoir des Word, Excel, des photos d'accidents PNG/JPEG dans les dossiers juridiques et des e-mails au format MSG.	Le format le plus courant est le PDF. Il pourrait y avoir des images Jpeg pour certaines copies d'écran mais cela ne devrait pas être le cas car elles sont mises en PDF.
3.3. Fichiers corrompus	Non.	Non.	Non.	Non.	Il y a eu des cas de notifications de fichier endommagé par l'ordinateur mais il était toujours possible de l'afficher et de le lire. Il n'y a pas eu de cas de fichiers disparus ou qui n'existent plus.	Non.

3.4. Volumétrie	Pas d'informations car cela est en cloud. Il y a un nombre total de 826 modules dont 113 sont archivés (inactif/formations passées).	Pas d'informations. Il y a une APOP par personne et par an dans la majorité des cas, avec les autres APOP intermédiaires il y a peut-être 2000 à 2500 APOP par an.	En termes de stockage Confluence ne représente pas grand-chose car il s'agit de texte. Le serveur de Confluence a peut-être 100 giga alors que probablement une centaine de personnes travaillent dessus.	Il y a environ 5GB de documents, cela représente la moitié de la capacité totale de l'espace de stockage. DiraJ	Il y aurait 6924 éléments au total depuis le début de GestE en 2014 (statut clos et statut en cours de tous les évènements qui ont un dossier juridique). En 2022, le total des fichiers sur GestE représente 209 GB.	La volumétrie du logiciel ou des éléments à archiver ne serait pas possible à quantifier, mais elle serait "énorme", il s'agit de plusieurs gigas. "Il y a plus de 300 écrans, Hastus couvre un très grand périmètre, ce sont des milliers de données." Il n'y a cependant pas de cas de surcharge, tout est monitoré et surveillé, il y a des back-ups.
3.5. Conservation des éléments historiques	Les modules de formations peuvent être "archivés" sur 360 Learning (ils sont alors inactifs), cela pourrait dépendre du choix du formateur. Voir aussi réponse 2.6.	Ne s'applique pas.	Interviewé 5 : Les éléments listés comme archives historiques seraient sous forme Wiki et non pas en pièces-jointes. Interviewé 1 : Le portefeuille de projet d'entreprise est dans la partie "Gestion de projet" de Confluence, il est géré par l'Exploitation. Tout ce qui est lié au Technique, par exemple, le Comité d'architecture technique est géré par l'IT.	Avec la somme des informations des documents listés, il n'est pas très clair et sûr de retrouver exactement ce qui est cherché.	Pour avoir plus d'informations sur les dossiers listés sous Exploitation (E), il faudrait contacter F. Fortin (Responsable du Service chantier et manifestations). Pour plus de détail sur les statistiques d'Exploitation du réseau TAC, il faudrait contacter R. Mary qui utilise QlikSense.	<i>Les informations concernant les éléments listés comme archives historiques devraient suffire pour retrouver chacun d'eux. Note (1) : Réf. SIT ne comprends pas pourquoi les documents listés devraient être archivés « puisqu'ils sont à tout moment sur Hastus même après la DUL ». Les documents sur Hastus remonteraient à 2009.</i>

						<i>Note (2) : Après relancement par e-mail : Les données entre 1998 et 2009 ne sont plus accessibles.</i>
3.6. Formats atypiques	Non.	Non.	Non. Au format Wiki (texte) en majorité.	Non.	Non.	Non.
4. Extraction						
4.1. Moyens d'extraction	<p>Pour exporter le contenu des formations outre les pièces-jointes, il n'y a pour l'instant pas de solutions. Il n'y aurait pas d'API.</p> <p>Cependant, il est possible de créer des formations dans un format standardisé SCORM qui pourrait ensuite être extrait en ZIP et fonctionner hors-ligne avec des pages HTML.</p>	<p>Il serait possible d'extraire le contenu des tables de données des APOP sur Allegro (avec un affichage sous forme de tables), ceci à l'aide de requêtes SQL par le SIT.</p> <p>Si l'on souhaite extraire les APOP au format RTF automatiquement, il n'y aurait pas d'API, il faudrait faire une demande à Adequasys (le fournisseur) pour nous les extraire dans le format souhaité.</p>	<p>Interviewé 1 : <i>Il est possible de télécharger manuellement un document à partir de Confluence.</i></p> <p>Interviewé 5 : <i>Il serait possible d'extraire automatiquement les éléments listés soit avec un Plug-in ou soit avec une API.</i></p>	<p>Il n'y a pas d'API. Il faudrait passer par le prestataire pour lui demander un script pour extraire tous les documents. Il faudrait voir comment est-ce qu'on nomme ces documents. Se demander si on garde leur nom original, il faudrait indiquer un numéro de dossier.</p>	<p>Pour extraire les éléments souhaités, il faudrait voir avec le prestataire (fournisseur de GestE). Une fonction existe pour imprimer manuellement en PDF depuis GestE.</p> <p>L'impression d'un évènement (dossier juridique) génère dans un PDF : les parties prenantes, les échanges, les photos etc. Mais il faudrait appliquer des droits spécifiques car seul le Service juridique devrait avoir le droit de le faire.</p>	<p>Il n'y a pas d'API pour Hastus, il faudrait utiliser les outils d'extraction manuelle présents sur Hastus avec l'aide des référents métier. Il faudrait prendre des rendez-vous avec le métier et que celui-ci lance les extractions.</p>

4.2. Moyens concrets	Pas d'informations.	Ne s'applique pas.	Dans le cas d'une demande plus approfondie d'extraction il faudrait revoir directement avec l'interviewé 5.	Pas d'informations.	Ne s'applique pas.	Ne s'applique pas.
4.3. Difficultés prévisibles	Voir réponse 4.1.	Si ce sont les tables de données des APOP qui sont extraites, les textes écrits lors de l'évaluation se retrouveraient dans des tables séparées.	Il n'y aurait en principe pas de soucis particuliers.	Pas d'informations.	Non, cela dépendrait de la manière dont l'on souhaite extraire.	Interviewé 12 (SIT), recommande de réfléchir à la plus-value de l'extraction des éléments souhaités car il s'agirait de tâches rébarbatives. Cette action devrait être argumentée de manière précise.

Partie 2

	Site Intranet	Isilog	MailChimp	QlikSense	SAP
1. Profil et description					
1.1. Réf. Métier : Présentation	Interviewé 13 (MVC), Chargé de communication interne. Fait en sorte que les messages les plus importants de l'entreprise soient transmis au personnel. Coordonne les magazines internes "Feu vert" et "Le reflet". Travaille avec les RH autour de la campagne contre le harcèlement et avec la DG lors de la crise Covid.	Interviewé 1 (DG), Gestionnaire des risques d'entreprise et Responsable du Management d'entreprise. Fait partie d'une équipe qui aborde les thématiques SQSE (santé, sécurité, sûreté, qualité), gestion du portefeuille de projet, la gestion documentaire et l'environnement.	Interviewé 15 (DG), Porte-parole et responsable des affaires publiques. Prépare et diffuse les communiqués de presse. Fait le travail de communication avec l'Etat de Genève, suit le travail du parlement et fait un travail de communication vis-à-vis des parlementaires.	Interviewé 11, E, Responsable Pôle Conception de l'offre.	Interviewé 16 (F&G) est Responsable Comptabilité/Fiscalité. L'activité principale consiste à produire les États financiers consolidés du groupe tpg et toutes ses filiales une fois par an. Pour y arriver, il faut tous les jours enregistrer les pièces comptables qui arrivent où qui doivent sortir et vérifier les imputations, la TVA et tout aille bien d'un point de vue comptable et fiscal.
1.2. Réf. Métier : Tâches en lien avec l'app. métier	Mise en ligne sur l'intranet les différentes communications qui parviennent des DA, alimente régulièrement la page d'accueil et crée certaines courtes actualités sur le site intranet.	Tous les rapports d'audit sont tenus en compte et les non-conformités ou recommandations sont entrées dans le système pour être suivies.	L'infolettre (ou newsletter) est un outil parmi d'autres en ce qui concerne la communication. Il s'agit de faire un "Best of" trimestriels de l'actualité qui touche l'entreprise et destiné au grand public.	Réalise des statistiques.	Toutes les tâches mentionnées en 1.1. se font sur SAP.
1.3. réf. Métier : Fréquence d'utilisation	Tous les jours. (Remarque : intranet était très peu utilisé et pauvre en contenu, il a été repensé lors de l'arrivée de l'interviewé 13 il y a 3 ans).	Pas d'informations.	Utilisation de MailChimp pour la création et l'envoi d'une newsletter une fois tous les 3 mois (trimestriel).	Tous les jours.	Tous les jours.

1.1. réf. SIT : Présentation	Interviewé 14 (SIT), Chef de projet IT.	Interviewé 14 (SIT), Chef de projet IT.	Interviewé 8 (SIT), Chef de projet IT. Gère le cycle de vie logiciel, répond aux besoins de solutions et teste les nouvelles solutions.	Pas de référent SIT interrogé.	Interviewé 17, SIT, Chef de projet IT. Travaille à mi-temps sur SAP finance et contrôle de gestions et à mi-temps sur la partie du système de vente.
1.2. réf. SIT : Tâches en lien avec app.	Assure la maintenance de l'application, les mises à jour, les évolutions vis-à-vis du métier et le lien avec le sous-traitant et le fournisseur. Fait en sorte que les demandes du métier concernant l'application soient justifiées et réalisables.	Gère le cycle de vie d'Isilog mais celui-ci va être décommissionné.	MailChimp n'est pas géré et il n'y a pas d'influence sur l'évolution de ce logiciel.	Pas d'informations.	Il y a différentes compétences sur SAP : finance, contrôle et gestion (Interviewé 17), administration, ventes, achats, RH, gestion des stocks.
1.3. réf. SIT : Fréquence	La maintenance des mises à jour ne prend pas énormément de temps dans la mesure où c'est le fournisseur qui les fait.	Seulement de manière ponctuelle.	Voir réponse 1.2.	Pas d'informations.	Tous les jours.
1.4. Type d'app.	Site intranet.	Isilog est un gestionnaire de suivi d'audits.	Solution de newsletter et une plateforme de marketing intégrée. MailChimp est une solution web de service à la demande, le logiciel n'est pas installé au tpg mais fonctionne en ligne.	QlikSense est un logiciel d'analytique, il fait de l'analyse de données. Il s'agit d'un client léger en local tpg.	SAP est un ERP (Enterprise Resource Planning).

1.5. Fonction app.	Intranet est l'outil principal pour diffuser la communication interne.	Isilog est l'outil de suivi des fiches d'amélioration (FA) d'entreprise. Permet de faire le suivi des audits.	MailChimp permet d'accompagner la mise en page des infolettres et de voir des statistiques liées au centre d'intérêt des destinataires (par exemple : les informations recherchées et là où les lecteurs cliquent).	QlikSense permet de faire des statistiques.	SAP est un progiciel intégré qui gère tout : les processus achats, les stocks, les ressources humaines etc. A la fin tout se déverse dans la comptabilité
1.6. Ancienneté app.	L'interviewé 14 précise qu'il y a eu un ancien site intranet qui a été complètement revu et redessiné en 2017 ou 2018. Avant cela il y avait un autre intranet qui était basé sur une autre technologie et qui existait depuis 2008 ou 2009. (L'interviewé 13 conseille de consulter J. S. pour plus d'informations).	Existerait depuis 2008. La version actuelle d'Isilog (2.7.7 IWS 2009) sera bientôt décommissionnée car elle n'est plus maintenue par son fournisseur (Isiline). Il n'y a plus eu de mises à jour depuis 2009. L'outil ne s'ouvre que dans Internet Explorer (navigateur précédant Microsoft Edge).	Aux tpg, la première infolettre avec MailChimp remonte à septembre 2018.	QlikSense existe depuis 2015 aux tpg. Il y a des mises à jour régulières tous les ans.	SAP a été mis en place aux tpg en 1996. La prochaine version de SAP est prévue pour 2025, elle se nomme S/4.

2. Cycle de vie					
2.1. Parcours/phases de vie	L'essentiel des documents sont faits par les domaines d'activité (DA) ou le management, ils rédigent, finalisent et valident les informations. Les documents (PDF) sont ensuite envoyés à l'interviewé 13 qui les met en ligne sur le site intranet. Par initiative personnelle, l'interviewé 13 conserve depuis 1 an les documents à mettre en ligne dans un espace personnel de l'ordinateur. Concernant les actualités en ligne faisant partie de l'affichage du site, elles sont toutes disponibles et conservées indéfiniment dans l'historique du site.	Des fiches sont créées sur Isilog et des rapports d'audits sont joints à la fiche. Lors de la création d'une fiche plusieurs métadonnées sont rentrées. Un responsable va consulter la fiche et la valider. Tous les rapports d'audit sont dans eDocs, Isilog n'a pas pour vocation de sauver des documents probants.	L'infolettre est créée sur MailChimp par le Porte-parole et responsable des affaires publiques. Elle est validée par le directeur général et envoyée. La newsletter est conservée dans la plateforme en ligne de MailChimp et dans les boîtes E-mail d'Outlook au format MSG. L'infolettre peut être convertie en format PDF. Elles sont conservées indéfiniment.	Les statistiques sont créées sur QlikSense puis exportables. Il est possible de faire de faire des revues de performance, des Data Story et des présentations de type PowerPoint.	D'un point de vue règles comptables, on ne peut pas avoir une pièce comptable créée dans SAP sans une pièce comptable en format papier. Les pièces comptables sont numérisées en PDF et ne sont pas modifiables. La pièce originale en papier est conservée dans des boîtes d'archives.
2.2. Assignment statut manuel	Ne s'applique pas.	Voir réponse 2.1. et 2.3.	Non/Ne s'applique pas.	Ne s'applique pas.	Voir réponse 2.3.
2.3. Assignment statut automatique	Ne s'applique pas.	Le statut de la fiche passe automatiquement de « demande d'attribution » à « validée », « suivie » et finalement « clôturée ».	Non/Ne s'applique pas.	Ne s'applique pas.	SAP indique automatiquement dans le statut du document si une facture est conforme ou pas (0 – facture pas encore traitée, 1 – traitée sans erreur, 2 – traitement erroné sur la facture). La comptabilisation de la pièce est cependant manuelle.

2.4. Assignment des durées de conservation	Ne s'applique pas.	Non/Ne s'applique pas	Non/Ne s'applique pas.	Ne s'applique pas.	Il n'y a pas d'évaluation sur SAP, les pièces comptables sont gardées indéfiniment tandis que les versions papier sont supprimées après 10 ans. Avec S/4, des règles d'archivage devraient permettre d'assigner des durées de conservation.
2.5. Élimination/transfert automatique/interconnexions	Ne s'applique pas.	Non.	Non/Ne s'applique pas.	Non.	SAP n'est pas très pratique pour les exports et n'est pas très user-friendly.
2.6. Pratiques version finale	Les documents PDF présents sur le site sont remplacés lors de la nouvelle version. Les actualités passées ne sont pas supprimées mais peuvent être modifiées pour indiquer qu'elles sont résolues.	Conservation.	Infolettre conservée sur MailChimp et au format MSG sur Outlook. Pourrait être convertie au format PDF.	Les statistiques ne sont pas forcément exportées en PDF (pas de manière systématique). Les revues de performance sont gardées tels quelles sur QlikSense. Le reste des informations reste sur QlikSense.	Voir réponse 2.4. Les pièces comptables numériques sont conservées indéfiniment dans la version actuelle de SAP.
2.7. Difficultés de recherche	Pas de difficultés mais certains collaborateurs ont parfois des difficultés à se retrouver sur intranet.	Non, la recherche est très simple.	Non, recherche très simple et intuitive.	Non, il serait possible de retrouver les statistiques souhaitées facilement.	Non.

2.8. Renseignement des métadonnées	Non, les documents arrivent au format PDF et il n'y a pas de modifications ou d'ajout aux métadonnées.	Plusieurs métadonnées sont rentrées : la source de la demande (audit interne ou externe), le responsable de l'action, le responsable SQSE (SQSE = Santé, Qualité, Sécurité, Environnement) etc. Toutes les métadonnées demandées sur Isilog avec des « * » rouges doivent être renseignées obligatoirement, Isilog veille à ce que ces champs ne restent pas vides.	Il n'y a pas de métadonnées renseignées, ni dans le cas d'une conversion au format PDF.	Non.	Il y a des champs obligatoires : la date de pièce, la date comptable, la devise, le type de pièce (s'il s'agit d'une facture fournisseur, un client, un débit ou un crédit), la clé de comptabilisation, la nature comptable, le code TVA, le montant et l'information analytique (une information d'un ordre ou d'un centre pour la comptabilité analytique). Ces champs sont toujours remplis, sinon ce n'est pas possible de passer à la page suivante car SAP bloquerait.
2.9. Gestion des données sensibles	Le référentiel classe le site intranet des tpg comme S1 (données personnelles non-sensibles).	Le référentiel classe Isilog comme S0 (données sensibles/confidentielles). Il y a des audits confidentiels.	Il s'agit d'information publique sélectionnée (S3).	Le référentiel classe QlikSense comme S2 (données internes).	Le référentiel classe SAP comme S0 (données confidentielles).
3. Conservation					
3.1. Garantie de valeur probante	Non/Ne s'applique pas.	Non. Il y a un horodatage à la création de la fiche mais tout est modifiable.	Non.	Non/Ne s'applique pas.	Non/il n'y a pas d'horodatage etc.

3.2. Formats	Les DA envoient les fichiers déjà prêts en format PDF. Il n'y aurait pas d'autres formats pour les documents mis en ligne sur l'intranet.	Il est possible d'ajouter des pièces-jointes aux fiches : Word, Excel, PowerPoint, PDF etc. (Suite Microsoft Office).	Pas d'informations. Les infolettres envoyées par MailChimp sont conservées sur Outlook au format MSG.	L'affichage sur QlikSense n'est pas dans un format particulier, c'est à l'export que le contenu acquiert un format, le contenu est exportable en PDF. (Ou aussi en Excel ou format image selon G. C.)	Les extractions se font le plus souvent sous Excel ou bien en CSV, PDF et il est possible de générer des formats HTML. Il s'agirait uniquement de PDF simple et non de PDF/A.
3.3. Fichiers corrompus	Non.	Non.	Non.	Non.	Non.
3.4. Volumétrie	L'interviewé 14 n'a pas d'informations par rapport à la volumétrie du site intranet. "Il faudrait regarder sur le serveur qui contient l'intranet pour avoir des données exactes."	Isilog doit peut-être représenter une centaine de mégas.	Il y a 12 infolettres à ce jour (en mai 2022).	Selon G. C. cela peut représenter plusieurs gigas selon le projet voir des centaines de méga.	Pas d'informations. Il n'y a pas de problèmes de capacité. "Le nombre de pièces comptables en 2021 était d'un peu plus de 1,7 millions d'écritures." Il faudrait préciser ce qu'on entend par "Plan comptable" dans la liste des éléments à archiver. <i>Note C. S. : Il s'agit uniquement du plan de base, pas question de tout garder.</i>

3.5. Conservation des éléments historiques	Tous les documents listés sont sur intranet et les informations/descriptions renseignées sur la liste des archives historiques seraient suffisantes pour les retrouver. Il faudrait préciser si la « politique du personnel » correspond au statut du personnel. Tous les documents importants sont sur la rubrique Kiosque.	Tous les documents listés comme historiques se retrouveraient bien sur Isilog.	Voir réponse 2.1.	Les statistiques remonteraient à 2015.	Pour la partie des documents listés en RH, il faudrait voir avec Y. M. ou avec l'interviewé 4 du côté RH ou bien E. R.
3.6. Formats atypiques	Non.	Non.	Non.	Non.	Non.
4. Extraction					
4.1. Moyens d'extraction	Il faudrait aller sur le serveur <i>Joomla!</i> (CMS) qui contient le site intranet et aller dans le répertoire des pièces-jointes pour tout extraire. Le nom des fichiers peut ne pas correspondre aux titres des documents. Pour effectuer des captures de l'aperçu du site intranet, c'est un outil du commerce qui doit permettre de scanner le site. Pour savoir jusqu'à quand remontent les back-ups du serveur il faudrait voir avec I&O. L'intranet ne contiendrait pas de Flash ni de JavaScript.	Il n'y a pas de moyen de télécharger toutes les fiches automatiquement. Les fiches sont dans une base de données en plus des fichiers qui sont sur le serveur. Pour extraire toutes les fiches, il est possible de faire une requête SQL. La base de données représente sous forme de tables le contenu de ces fiches. Pour les fichiers joints, il faudrait aller sur le serveur afin de les télécharger.	Le plus simple serait de faire l'extraction manuellement vu le petit nombre d'infolettres.	G. C. conseille de faire un point des documents précis, pour savoir exactement ce que l'on veut et que les référents métier puissent les extraire manuellement depuis QlikSense, l'extraction peut être faite en PDF, image ou Excel.	Ne pense pas qu'une API soit utilisée. Il est cependant possible de <i>batcher</i> des programmes. La solution d'archivage devrait avoir des interfaces avec SAP.

4.2. Moyens concrets	Ne s'applique pas.	Isilog sera décommissionné et la question de la migration des données va être abordée. Il y aura un travail qui sera fait par le fournisseur (Isiline) si on sélectionne ce fournisseur pour reprendre la suite d'Isilog	Pas d'informations.	Pas d'informations.	Pas d'informations.
4.3. Difficultés prévisibles	Ne s'applique pas.	Non.	Durant la démonstration en fin d'entretien, des problèmes d'affichage ont surgi lors de la conversion au format PDF d'une infolettre sur Outlook. La conversion PDF à partir de l'utilitaire d'impression du client web (navigateur internet) s'est avérée plus fidèle. Les infolettres contiennent des liens vers le site des tpg et aussi des sites externes à l'entreprise. Les liens vers ces sites sont cliquables sur le PDF.	Non.	Il ne devrait pas y avoir de problèmes à convertir des Excel en PDF dans le cas de SAP. Il faudra voir du côté RH si le fait qu'il y a des données sensibles pose problème, et que ce ne soit accessible qu'aux personnes qui en ont le droit.

Partie 3

	Site web internet	SmartCockpit	Solu-Qiq	Teams	eDocs
1. Profil et description					
1.1. Réf. Métier : Présentation	Interviewé 18, MVC, Webmaster à l'Information voyageur (IV) : S'occupe de la gestion des supports numériques (écrans embarqués et aux arrêts) et de la gestion du site web mais pour la partie IV uniquement. Interviewé 19, MVC, Coord. Info Voyageurs Terrain : Est Responsable à l'intérim de l'Information voyageur (IV). Génère à destination du voyageur en support tout ce qui est plan de réseau schématique et géographique que ce soit imprimé ou numérique.	Interviewé 1 (DG), Gestionnaire des risques d'entreprise et Responsable du Management d'entreprise. Fait partie d'une équipe qui aborde les thématiques SQSE (santé, sécurité, sûreté, qualité), gestion du portefeuille de projet, la gestion documentaire et l'environnement.	Interviewé 1 (DG), Gestionnaire des risques d'entreprise et Responsable du Management d'entreprise. Fait partie d'une équipe qui aborde les thématiques SQSE (santé, sécurité, sûreté, qualité), gestion du portefeuille de projet, la gestion documentaire et l'environnement.	Ne s'applique pas.	Interviewé 23 (réf. métier), Spécialiste en gestion documentaire. A pour mission de faire en sorte que la bonne information soit accessible aux bonnes personnes, au bon moment, dans les bonnes versions. L'objectif est de mettre à plat l'ensemble des processus de gestion documentaire aux tpg. Travaille surtout au niveau des processus, pour rationaliser les méthodologies de travail et mettre en place des règles de gouvernance documentaire.
1.2. Réf. Métier : Tâches en lien avec l'app. métier	Voir les tâches mentionnées en 1.1.	Pas d'informations.	Pas d'informations.	Ne s'applique pas.	Met dans la GED les documents aboutis ou en version finale. Gère tous ses documents dans eDocs sauf les documents en cours de création.
1.3. réf. Métier : Fréquence d'utilisation	Quelques fois par semaine dépendant des tâches (représente environ 15%-20% de l'activité).	Pas d'informations.	Pas d'informations.	Ne s'applique pas.	Tous les jours.

1.1. réf. SIT : Présentation	Interviewé 8 (SIT), Chef de projet IT. Gère le cycle de vie logiciel, répond aux besoins de solutions et teste les nouvelles solutions.	Interviewé 20 (SIT) a recommandé d'effectuer l'entretien avec l'interviewé 1 (DG).	Interviewé 21, SIT, Responsable Sécurité Informatique. Tâche principale : réduire les risques de l'entreprise en rapport avec l'information et des personnes concernées avec des mesures de sécurité.	Interviewé 22 (SIT), Responsable infrastructure et opération. Rôle d'assurer que les systèmes informatiques fonctionnent au quotidien. Était responsable des plateformes : GED, 360Learning et TEAMS jusqu'en 2021.	Interviewé 22 (SIT), Responsable infrastructure et opération. Rôle d'assurer que les systèmes informatiques fonctionnent au quotidien. Était responsable des plateformes : eDocs, 360Learning et TEAMS jusqu'en 2021.
1.2. réf. SIT : Tâches en lien avec app.	Gère le cycle de vie du site internet.	Pas de référent SIT interrogé.	L'unité s'occupe cartographier le système d'information. Le but est de montrer clairement les interfaces entre les différents systèmes et différents services (de manière visuelle).	Gérait la partie technique de Teams jusqu'en 2021.	Gérait la partie technique de eDocs jusqu'en 2021.
1.3. réf. SIT : Fréquence	N'influe pas sur le contenu du site internet.	Pas de référent SIT interrogé.	Pas précisé.	Tous les jours.	Rarement.
1.4. Type d'app.	Site web internet.	Smart Cockpit est un gestionnaire de données, un gestionnaire d'indicateurs. Il s'agit d'un client léger. Smart Cockpit est aussi un outil du système de management intégré, c'est l'outil de suivi des indicateurs utilisé par le collège de direction.	Solu-Qiq est le manuel qualité d'entreprise. L'outil contient l'organisation de l'entreprise (sous forme de cartographies, organigrammes). Il rend possible d'explorer les macro-processus sous forme de cartographie cliquable sur plusieurs niveaux.	Logiciel collaboratif.	GED (Gestion électronique des documents).

1.5. Fonction app.	Fonctions bien connues du site web des tpg (Le site internet tpg réunit toutes les informations relatives aux tpg destinées au public, clients et aux voyageurs).	SmartCockpit permet de faire des tableaux de bord (cockpits). L'outil permet d'avoir une vue d'ensemble en temps réel de la situation de l'entreprise à l'aide d'indicateurs sur l'économie, les risques etc.	Rendre accessible l'information concernant les processus, organigrammes de l'entreprise etc.	Teams est avant tout un outil de communication, pour passer des appels audios et des visioconférences. Les collaborateurs partagent et travaillent sur des documents à court terme. La question de la conservation ne se pose pas.	Comment le fait que la GED n'est pas utilisée à son plein potentiel et que cet outil, faute de règles, est utilisé différemment par chaque collaborateur comme un outil collaboratif ou comme un stockage de plus. Seul les versions abouties, formalisées ou les versions finales des documents devraient se retrouver dans eDocs.
1.6. Ancienneté app.	La version actuelle du site "WEX" date de décembre 2019. Historiquement, le site existerait depuis 2007. L'interviewé 18 précise qu'il y a eu deux changements de version : en 2011 et 2019. Wayback Machine retrace cependant le site des tpg jusqu'à janvier 1999.	Existerait depuis 2016.	Existe aux tpg depuis 2013-2014.	Sortie de Teams en 2017 dans la suite Microsoft 365.	La GED eDocs est ancienne, elle a plus de 20 ou 30 ans.

2. Cycle de vie					
2.1. Parcours/phases de vie	<p>Les Rapports de gestion et les Communiqués de presse sont rédigés par la DG et mis en ligne par la Communication (avec comme responsable : J. S.). Pour les questions concernant les Communiqués de presse et le Rapport de gestion voir avec I. P. (Communication). Sur le site web, le plus ancien communiqué de presse sur le site web remonte au 24 février 2011 et le plus ancien Rapport de gestion à 2016 (en PDF).</p>	<p>Les données présentes sur SmartCockpit sont alimentées en temps réel, ces données sont issues d'autres outils, SmartCockpit est une agrégation de ces données qui permet d'avoir une vue large de l'entreprise. Il est possible de retrouver des fichiers joints aux indicateurs. SmartCockpit est composé d'un cadran qui contient la liste de indicateurs, les indicateurs contiennent le détail des données. Le cadran et les indicateurs sont exportables au format Excel.</p>	<p>Un document, par exemple, une procédure d'exploitation peut soit être créée directement sur Solu-Qiq, soit être importée d'un fichier Word, Excel, PDF (ou du reste de la suite Microsoft Office). Il peut y avoir des fichiers joints aux cartographies. Les cartographies sont créées dans Solu-Qiq. P. C., Responsable environnemental, s'occupe de coder le tout, les DA viennent avec des schémas pour lui présenter leurs processus afin de les mettre sur Solu-Qiq.</p>	<p>Dès que le document a été travaillé/finalisé, il est retiré et mis ailleurs.</p>	<p>Création du document dans la GED ou dans un autre environnement (il n'y a pas de règles). Normalement, lorsqu'un document est dans sa version aboutie ou formalisée, il doit entrer dans la GED. Une fois que ce document est validé, la DUA est déclenchée. Plus tard, il passe en DUL puis vient la clôture. Tant qu'un document n'est pas clôturé, il va vivre, il va être partagé, il va être transféré, on va le consulter, etc. À noter qu'il n'y a pas de dossiers dans la GED, les documents sont uniquement virtuellement reliés entre eux et possèdent un identifiant.</p>
2.2. Assignment statut manuel	Ne s'applique pas.	Non/Ne s'applique pas.	Non/Ne s'applique pas.	Non.	Devrait être implémenté dans le cadre d'une étude d'analyse et développement.
2.3. Assignment statut automatique	Ne s'applique pas.	Non/Ne s'applique pas.	Non/Ne s'applique pas.	Non.	Devrait être implémenté dans le cadre d'une étude d'analyse et développement.

2.4. Assignment des durées de conservation	Ne s'applique pas.	Non/Ne s'applique pas.	Non/Ne s'applique pas.	Non.	Non, mais il est prévu dans le cadre d'une étude. Il faut définir exactement ce qui déclenche la DUL.
2.5. Élimination/transfert automatique/interconnexions	Non/Pas d'informations.	Non.	Non.	Non.	Pour un transfert, il faudrait que la solution de d'archivage et pérennisation long terme puisse identifier les aspects de eDocs.
2.6. Pratiques version finale	Les versions mises sur le site ont des copies dans des dossiers sur les serveurs tpg (où l'on retrouverait également des anciennes versions).	Si les indicateurs ne sont plus utiles, ils sont supprimés, ou en tout cas supprimés de l'affichage (leur créateur pourrait continuer de les voir mais pas les autres utilisateurs).	Lorsque, par exemple, une nouvelle version d'organigramme apparaît, l'ancienne est écrasée et remplacée par la nouvelle.	Dès que le document a été travaillé/finalisé, il est retiré et mis ailleurs.	Un document qui serait en version finale est gardé dans la GED. Si le document est arrivé à la fin de sa DUL et est clôturé, il devrait être éliminé ou conservé. En pratique les documents sont pour l'instant uniquement conservés.
2.7. Difficultés de recherche	Non.	Non.	Non, il faut toutefois savoir sous quel DA se trouve l'information pour aller la chercher.	Ne s'applique pas.	Oui, il y a eu des difficultés dans la recherche. Cependant, avec la version client léger (version web) prochainement mise en service, la recherche serait plus rapide et efficace. Il est possible de faire des recherches plein texte.

2.8. Renseignement des métadonnées	Non.	L'export en Excel ne contient pas de métadonnées telles que l'auteur ou le titre. Il est possible d'ajouter une petite description aux documents lorsqu'ils sont téléversés sur SmartCockpit.	Les documents importés peuvent être exportés avec les mêmes métadonnées que lors de l'importation. La démonstration de Solu-Qiq montre qu'en réalité il n'y a pas de métadonnées pertinentes renseignées lors d'un export PDF d'un visuel. Seul le titre du visuel tel qu'il est indiqué dans Solu-Qiq est renseigné. La date de création correspond faussement à la date d'export.	Non, pas possible sur Teams.	Réf. SIT : Il est possible de renseigner les métadonnées au niveau des documents. Un ancien formulaire de saisie existe mais il ne fait pas la part belle aux métadonnées, il faudrait un référentiel de métadonnées pour les rendre utilisables pour améliorer la recherche par d'autres personnes. Réf. métier : eDocs peut récupérer automatiquement le titre à l'intérieur d'un document pour le mettre dans le champ titre des métadonnées. Il y a le souhait de simplifier la saisie et de n'exiger que de simples métadonnées : Titre, date, auteur, type de document.
2.9. Gestion des données sensibles	Le référentiel classe le Site web internet comme S3 (données publiques).	Le référentiel classe SmartCockpit comme S2 (données internes : informations relatives au fonctionnement des services)	Le référentiel classe solu-Qiq comme S2 (données internes : données relatives au fonctionnement des services).	Teams est classé S2 (données internes) selon le Référentiel services et traitements.	eDocs est classé S2 (données internes) par le Référentiel services et traitements. Les documents devraient être classés sous un des trois éléments : public, interne ou confidentiel. Mais aucune réflexion n'a été menée pour définir le sens sémantique de ces éléments. Il faudrait plus

					de cohérence, la gestion du droit d'accès est au bon-vouloir de l'auteur du document.
3. Conservation					
3.1. Garantie de valeur probante	Non/Ne s'applique pas.	Non.	Non/Ne s'applique pas.	Non.	Oui du point de vue des exigences internes. Non du point de vue des exigences juridiques.
3.2. Formats	Les documents listés comme historiques seraient au format PDF. Le site contient plusieurs formats d'images : JPEG, PNG, PDF. Il y aurait des fichiers vidéo MP4.	Il est possible de joindre aux indicateurs des fichiers PDF et de la suite Microsoft Office.	Il y a des fichiers Excel, Word, PDF (et autres formats standards de la suite Microsoft). Ces fichiers ne sont pas reliés à la GED eDocs, il peut s'agir de copies de la GED eDocs mises dans l'outil.	Les plus courants sont Word, Excel et Powerpoint.	Tous les formats. Les plus courants sont Word, PowerPoint et PDF. Il peut y avoir des Excel.
3.3. Fichiers corrompus	Cela peut arriver lors d'évolutions mais en général il n'y a pas de problèmes.	Non.	Non.	Il n'y a jamais eu de cas de perte ou fichier corrompu reporté sur Teams.	Il y a eu des fichiers corrompus, des affichages déformés et fichiers qui n'existaient plus. Il y a eu des bugs de logiciel. Mais les problèmes ont été évités grâce aux nombreuses sauvegardes et restaurations.

3.4. Volumétrie	Les fichiers média sont conséquents. Pour plus d'informations, il faudrait voir avec les ingénieurs infra.	Il y a une cinquantaine de cockpits et environ 200 indicateurs (mais ils ne sont pas tous actifs)	La volumétrie de Solu-Qiq n'est pas très conséquente, il s'agit de descriptif, de textes, ce qui peut ne représenter que quelques mégas probablement.	Pas d'informations, ce qui est collectif est relativement léger et les espaces individuels ne peuvent pas être estimés.	Le volume de documents est très important dans la GED (307'000 fichiers), il n'y a pas de vision très juste de la façon de stocker (GED utilisée comme espace de stockage). Il y a eu 21'000 documents ajoutés en 2021, en moyenne 20'000 nouveaux documents sont ajoutés chaque année depuis 2015.
3.5. Conservation des éléments historiques	Les documents listés seraient tous au format PDF.	Voir réponse 2.1.	Le « Recueil de procédures comptables - Système de contrôle financier interne » marqué comme « à vérifier » dans la liste des archives historiques se trouverait aux finances (sous H. F., trésorière et gérante du Système de contrôle interne financier (SCI)).	Pas d'informations.	Pour savoir comment ils sont conservés, il faudrait préciser le numéro d'identifiant des documents historiques.
3.6. Formats atypiques	Il n'y aurait pas de formats atypiques.	Non.	Non.	Pas d'informations.	Voir réponse question 3.2. et 3.5.

4. Extraction					
	<p>Si l'on prend l'exemple des communiqués de presse, ils ne sont pas tous classés dans le même dossier ni nommés uniformément, ce qui rend leur recherche/extraction plus difficile. Une extraction manuelle serait à envisager, sinon, el faudrait voir si un autre moyen d'extraction est possible avec le prestataire du site web (ELCA).</p> <p><i>Note : Il n'y a pas de back-ups des versions du site internet avant 2012.</i></p>	<p>Pour une capture complète des données présentes sur SmartCockpit il faudrait une capture du cadran pour voir la liste de tous les indicateurs existants, puis effectuer un téléchargement à partir de chaque indicateur pour avoir le détail des données. Toutes ces informations seraient extraites en format Excel. Il s'agirait d'une tâche conséquente. On pourrait aussi imaginer de faire la demande d'extraction auprès de l'entreprise qui fournit SmartCockpit.</p>	<p>Il est possible d'imprimer (extraire) en PDF à partir de Solu-Qiq.</p> <p>Si l'on souhaite conserver les visualisations et qu'elles soient cliquables, il faudrait faire un « enregistrer sous » en HTML dans le navigateur et pour chaque page afin de pouvoir mettre en marche les liens. Il serait plus simple et recommandé d'utiliser un aspirateur de page web pour tout absorber. Il n'y a pas d'API disponible pour cela.</p>	<p>On ne peut pas extraire sur Teams car celui-ci n'est pas cadré. Ce n'est qu'un outil de stockage à court terme.</p>	<p>L'éditeur de la GED eDocs devrait fournir une API pour l'extraction mais la difficulté serait de retrouver les documents listés comme historiques parmi la masse d'informations de la GED.</p>
4.1. Moyens d'extraction					

4.2. Moyens concrets	Pas d'informations.	Ne s'applique pas.	Pas d'informations/voir réponse 4.1.	Voir réponse 4.1.	<i>"Il faut que d'un côté que l'éditeur propose cette API et que qqn vienne utiliser cette app et la consommer pour pouvoir extraire des données. Dans l'un comme dans l'autre je n'ai pas de réponse. La première c'est à l'éditeur de la fournir, la deuxième au fournisseur de la solution d'archivage qui est celui qui sait interroger l'API pour en récupérer le contenu et retourner des résultats proposés par l'API."</i>
4.3. Difficultés prévisibles	Pas d'informations.	L'extraction faite maison aux tpg représenterait une tâche conséquente. La lisibilité des données extraites sur Excel est difficile et s'apparente à celle d'un affichage de données brut sur CSV par exemple.	Non.	Il n'y a rien qui permette de piloter ce qui se passe et de le sauvegarder et standardiser.	Il y a un fort risque d'altération de la qualité de lecture des documents. Rien ne garantit que l'affichage soit fidèle à la version originale. Le problème concernerait surtout les documents de type Word et PowerPoint. Pour les PDF, la garantie est un peu plus élevée tant qu'il s'agit de versions normalisées.

Annexe 6 : Liste des métadonnées pour l'archivage aux tpg

N° de référence	Titre de la métadonnée	Définition	Obligatoire / facultatif O/F	À reprendre pour le versement Oui/Non	Type et taille du champ	Notes
MD26	Nom du producteur	Nom du service/unité à l'origine de la production de l'ensemble des documents (produits ou reçus). Ecrire le nom en entier, sans abréviation.	O	Oui	Alphanumérique 120 caractères	Spécifique à chaque application. Le nom du producteur est défini lors de la création de la base de données en collaboration avec l'archiviste/SGD ; il est ensuite rempli automatiquement (pour autant qu'il soit unique, sinon un menu déroulant est créé).
MD0	Numéro de dossier	Référence administrative du dossier. Selon les services, cela peut prendre la forme d'un n° dossier sinistre, d'un n° de facture, etc.	F	Oui	Alphanumérique 70 caractères	L'intitulé du champ peut changer selon le métier (exemple: N. dossier sinistre; N. de facture...)
MD1 à MD19	Information d'identification supplémentaire du dossier	Caractéristiques supplémentaires permettant d'identifier le dossier et propres à chaque service.	F	Oui	Alphanumérique 120 caractères	Informations métier. Une de ces métadonnées spécifiques doit donner l'information de déclenchement du processus permettant de calculer la DUA/DUL (p. ex. : passage RD, fin d'année civile, clôture du dossier, fin d'année comptable, etc.)
MD22	Intitulé du dossier	Titre du dossier ou courte description de son contenu	O	Oui	Alphanumérique 120 caractères	
MD21	Rubrique du plan de classement	Intitulé de la rubrique du plan de classement (ajouter	O	Oui	Alphanumérique 120 caractères	

N° de référence	Titre de la métadonnée	Définition	Obligatoire / facultatif O/F	À reprendre pour le versement Oui/Non	Type et taille du champ	Notes
		éventuellement le numéro de position de la rubrique)				
MD23	État du cycle de vie	Etat du cycle de vie du document (menu déroulant)	O	Non	Liste déroulante (Document actif / Document intermédiaire / Archive définitive)	Le passage d'un stade à l'autre se fait automatiquement selon les durées entrées en DUA/DUL (à voir selon les capacités du système GED)
MD24	Type de document	Typologie (menu déroulant selon les types de documents répertoriés dans le chapitre 3 des règles de nommage) ou automatique s'il s'agit d'un modèle	O	Oui	Liste déroulante typologie de document	Automatique si issu d'un modèle
MD51	Sous-type de document	Typologie précise liée directement au calendrier de conservation qui permet de définir le MD51, MD35, MD36, MD39.	O	Oui	Liste déroulante précise	Exemple : Si MD24 = PV, MD51 peut être « de RD », « de CA », ou encore « d'équipe ». Spécifique tpg
MD25	Numéro de règle du calendrier de conservation	Numéro de la règle qui figure dans le calendrier de conservation	O	Non	Alphanumérique 120 caractères	Automatique si issu d'un modèle Automatique parce que issu sous-type MD25
MD27	Intitulé du document	Titre principal ou "Concerne" du document	O	Oui	Alphanumérique 120 caractères	Utilisation des règles de nommage
MD28	Auteur du document	La ou les personnes responsables du contenu du document. On considère comme auteur celui qui signe, celui qui assume la responsabilité du contenu du document.	F	Oui	Alphanumérique 120 caractères	Automatique mais modifiable

N° de référence	Titre de la métadonnée	Définition	Obligatoire / facultatif O/F	À reprendre pour le versement Oui/Non	Type et taille du champ	Notes
MD31	Editeur du document	Nom du ou des service(s), groupe(s), commission(s), etc. à l'origine de la publication du document (pas de noms de personnes!).	F	Non	Alphanumérique 120 caractères	
MD32	Contributeur du document	Nom d'une personne, du service, du groupe, etc. qui contribue ou a contribué à l'élaboration d'un document rédigé par des tiers.	F	Non	Alphanumérique 120 caractères	
MD29	Sujet du document	Mots-clés (noms communs, noms de lieu, noms propres), phrases de résumé.	F	Non	Alphanumérique 120 caractères	Besoin d'un langage contrôlé, c.à.d. d'un thésaurus. Peut être utile pour la recherche de documents.
MD30	Description du document	Description ou résumé du contenu du document sous forme rédigée.	F	Non	Alphanumérique 120 caractères	
MD33	Date de création du document	Date à laquelle le document est créé.	O	Oui	AAAA-MM-JJ	Automatique, par défaut date du jour mais modifiable. Actuellement « Date Doc » dans la GED.
MD34	DUA	Durée d'utilité administrative du document	O	Non	Numérique 120 caractères	Automatique si issu d'un modèle. Une des métadonnées spécifiques aux services MD3 à MD20 permet de déclencher le processus du calcul de la DUA/DUL.
MD35	DUL	Durée d'utilité légale du document	O	Non	Numérique 120 caractères	Automatique si issu d'un modèle
MD36	Sort final	Sort final du document (menu déroulant)	O	Non	Liste déroulante:	Automatique parce que issu du MD25.

N° de référence	Titre de la métadonnée	Définition	Obligatoire / facultatif O/F	À reprendre pour le versement Oui/Non	Type et taille du champ	Notes
					(Conservation / Echantillonnage / Destruction)	
MD39	Numéro unique	Numéro donné par l'application à chaque document dans la base.	O	Oui	Numérique 120 caractères	Automatique et correspondant au n° GED

Annexe 7 : Tableau simplifié PAIMAS

1) Phase préliminaire

<i>Sous-phases</i>	<i>Étapes</i>	<i>Sous-étapes</i>	<i>Description</i>	<i>Qui ? P=Producteur A=Archiviste</i>
PREMIER CONTACT		1. Identification des personnes et définition des modalités de collaboration Archivistes : C. Schneider et J. Collet Spécialiste en gestion documentaire : Interviewé 23 Producteurs : - GED eDocs : Interviewé 23 (réf. Métier et Spécialiste en gestion documentaire), Interviewé 22 (ancien référent SIT de eDocs) - Isilog : Interviewé 1 (réf. métier), Interviewé 14 (réf. SIT) - Disque Q : Contacts à définir (Producteurs : T, MVC et I&B)		P et A
		2. Echange d'informations générales		P et A
DÉFINITION PRÉLIMINAIRE, FAISABILITÉ, ÉVALUATION	<i>Information à archiver</i>	3. Identification du contenu d'information à archiver Dossiers et documents numériques à valeur historique cités dans la liste ad-hoc en annexe.		P et A
		4. Identification de l'information complémentaire (information de pérennisation) Métadonnées nécessaires, voir la liste des métadonnées pour l'archivage électronique : - Nom du producteur ; - Intitulé ; - Rubrique du plan de classement ; - Type de document ; - Sous-type de document ; - Intitulé du document ; - Date de création du document ; - Numéro unique. Ou au minimum : le titre et la date.	- sur la base d'une évaluation, inventaire des métadonnées nécessaires à la compréhension du contexte de création des documents et des données	P et A
		5. Identification de la communauté cible - GED eDocs : Données internes en principe, accès à étudier au cas par cas. - Isilog : Certaines données sont confidentielles, seuls les ayant droit définis par les responsables d'Isilog y ont accès. - Disque Q : Accès aux utilisateurs internes	- qui et par qui sont utilisés les documents et les données (utilisateurs internes et/ou externes)	P et A

		6. Définition des modalités d'accès aux documents et données Accès gratuit et total aux utilisateurs définis au point 5.	- modalités possibles (accès limité ou total, accès gratuit ou payant, niveau de service (rapidité, performance, type d'accès (par Internet, par réseau...), volumétrie...))	P et/ou A
		7. Définition du cycle de vie des documents et des données à archiver Conservation définie en détail dans le calendrier de conservation des tpg et aperçu du sort final dans la liste des archives à valeur historique en annexe.	- conservation temporaire et/ou définitive	P et/ou A
		8. Evaluation de la faisabilité et des coûts associés à cette sous-étape À déterminer entre l'archiviste et les producteurs.		P et/ou A
	<i>Objets numériques et standards appliqués à ces objets</i>	9. Identification préliminaire de la typologie des documents et des données à archiver Détails des documents et données à archiver dans la liste des archives à valeur historique en annexe. - GED eDocs : PV de séances, rapports divers, correspondance, finances, manuels, règlements. - Isilog : Manuel qualité, rapports, fiches d'amélioration, carte stratégique. - Disque Q : autorisations de construction, plans de véhicules, séances OPS, notes d'informations.		P et A
		10. Définition des normes et standards Formats pérennes et métadonnées selon les recommandations indiquées dans ce travail.	- formats de fichiers pérennes - modèles de métadonnées (information de pérennisation (PREMIS, METS...) et information de description (DC, EAD, MODS, MARCXML...))	A
		14. Evaluation de la compatibilité et étude des solutions		P et A
		15. Evaluation des efforts à accomplir et des coûts associés à cette sous-étape		P et A
	<i>Référence des objets</i>	16. Inventaire des informations sur les règles d'identification existantes et les nomenclatures		P et A
		17. Définition des règles d'identification		P et A
		18. Evaluation des coûts associés à cette sous-étape		P
	<i>Quantification</i>	19. Estimation du volume des documents et des données à transmettre aux Archives GED eDocs : Pas d'informations (181 documents présents sur eDocs sont listés pour l'archivage). Isilog : « Une centaine de méga » (Interviewé 14). Disque Q : - Q/Division/Doc_Entr/NI : 1,68 Go (5'755 fichiers, 24 dossiers). - Q/Division/C/SCO : 370 Ko (27 fichiers, 8 dossiers, non-accessible). - Q/Division/T/DAO : 9,59 Go (19'628 fichiers, 1534 dossiers).	- volume global, taille maximale des objets numériques, nombre de fichiers - estimation de la fréquence des versements	P

		- Q/Division/DRI :Bâtiments IB : non-accessible, vide ?		
		20. Evaluation du volume des documents et des données à pérenniser		A
		21. Evaluation de la capacité de l'infrastructure à recevoir les documents et les données (processus d'ingest)		A
		22. Evaluation des coûts associés à cette sous-étape		A
	Conditions de sécurité	23. Identification des prérequis nécessaires à la maintenance de la confidentialité des informations et à l'authentification de la source des documents et des données durant le processus de versement	- confidentialité : cryptage des documents et des données - authenticité : sommes de contrôle, signatures	P et A
		24. Identification des prérequis nécessaires à la sécurité des données au sein des Archives		A
		25. Identification des prérequis nécessaires à la maintenance de la confidentialité des informations et à l'authentification de la source des documents et des données durant le processus de diffusion de ces derniers		P et A
		26. Identification des normes et standards et des outils à utiliser - Téléchargement des documents et/ou dossiers dans un dossier de transfert. - Conversion des fichiers en format pérenne selon les recommandations. - Création du SIP avec Docuteam Packer. - Versement dans Docuteam cosmos.		P et A
		27. Evaluation des coûts associés à cette sous-étape		P et A
	Aspects légaux et contractuels	28. Définition de la nature des relations entre le producteur et les Archives Le producteur est tenu de préparer d'appliquer les métadonnées minimales et d'extraire les éléments demandés avec l'aide du SIT.		P et A
		29. Evaluation du problème de la propriété intellectuelle		P et A
		30. Définition des conditions d'accès aux documents et aux données GED eDocs : S2 (Données internes) Isilog : S0 (Données sensibles/confidentielles) Disque Q : pas d'informations, un dossier est à accès restreint.	- modalités d'accès durant la DUAL - modalités d'accès durant le délai de protection	P et A
		32. Identification des normes et standards et des outils à utiliser		P et A
	Opérations de transfert	33. Evaluation des coûts associés à cette sous-étape		P et A
		34. Définition préliminaire du SIP		P et A
		35. Echange des prérequis et des contraintes liés au respect des normes et standards de versement		P et A
	Validation	36. Evaluation des coûts associés à cette sous-étape		P et A
		37. Retour d'information au producteur sur les procédures de validation des SIP, les procédures de non-conformité et les outils utilisés par les Archives	- validation du modèle de SIP et du contenu des SIP	A
		38. Etude des développements ou des modifications à apporter aux outils de validation		A

		39. Etude de l'implémentation du contrôle qualité et des outils y relatifs		P
		40. Evaluation des coûts associés à cette sous-étape		P et A
	<i>Planification</i>	41. Définition d'une planification préliminaire	- planification du cycle de vie d'un versement	P et A
	<i>Impact permanent pour les Archives</i>	42. Evaluation de l'impact des procédures et des coûts associés sur les Archives <i>Contrôle du matériel, des logiciels et des formats de fichier.</i> <i>Opérer les migrations du matériel informatique nécessaires.</i> <i>Contrôle de l'intégrité des éléments versés.</i>	- volume permanent des documents et données à pérenniser (financement de l'infrastructure informatique) - actions nécessaires à la pérennisation (changements de l'infrastructure informatique, migrations des fichiers...) - mesures à prendre pour éviter une perte des documents et des données - mesures de sécurité	A
	<i>Résumé des coûts et des risques</i>	43. Evaluation des coûts et des risques liés à la création de ce nouveau processus de versement sur les court, moyen et long termes	- changements qui nécessitent de nouveaux investissements (nouveau flux, changements techniques...) - ressources humaines et financières	P et A
	<i>Points critiques</i>	44. Evaluation des points critiques		P et A
ÉTABLISSEMENT D'UN ACCORD PRÉLIMINAIRE		45. Elaboration d'un document récapitulatif	- document qui récapitule les analyses faites et les décisions prises durant cette phase	P et/ou A
		46. Elaboration d'un protocole préliminaire (en vue de la prochaine phase)	- contenu du SIP (données, information de pérennisation et information de description) et modèle de données - premier échéancier pour les versements - restrictions d'accès aux documents et aux données - clauses de révision et de renégociation du protocole préliminaire	P et A

2) Phase de définition formelle

Sous-phases	Etapes	Sous-étapes	Description	Qui
ORGANISATION DE LA PHASE DE DÉFINITION FORMELLE		1. Modalités d'organisation de la phase de définition formelle Phases : - Validation du choix des documents à archiver. - Repérage précis des documents dans l'applicatif métier ou dans le disque réseau. - Vérification et implémentation des métadonnées nécessaires. - Création des SIP avec Docuteam Packer. - Transfert vers Docuteam Cosmos. Documents produits par : - GED eDocs : DG, E, SIT, I&B, MVC, T, CA, F&G, RH (Tous les DA). - Isilog : DG, CA. - Disque Q : E, T, I&B, MVC.		P et/ou A
		2. Examen approfondi des points problématiques de la phase précédente		P et A
DÉFINITION FORMELLE	Information à préserver et modèle des objets de données à verser	3. Définition du contexte général du projet ainsi que de la liste et des contenus des éléments d'information à verser Se référer à la liste des archives à valeur historique.	- contenu d'information - information de pérennisation (provenance, contexte, référence, intégrité) - information de description	P et A
		4. Définition des formats de fichiers, des règles d'encodage et des normes et standards à appliquer aux objets d'information GED eDocs : DOC(X), XLS(X), PPT(X), PDF. Isilog : Contenu des fiches de la BDD Isilog (à convertir en SIARD), pièce-jointes actuelles en DOC(X), XLS(X), PPT(X), PDF. Disque Q : - Q/Division/Doc_Entr/NI : DOC, DOCX, XLS, PDF, TIFF. - Q/Division/C/SCO : XLSX, DOCX, DOC - Q/Division/T/DAO : DWG, DXF, IDW, PDF, JPEG, TIFF, XLS, XLSX, PPT, MSG, RTF, HTML, TXT. - Q/Division/DRI : Bâtiments IB : non-accessible.		P et A
		5. Définition de la volumétrie GED eDocs : Pas d'informations (181 documents présents sur eDocs sont listés pour l'archivage). Isilog : « Une centaine de méga » (Interviewé 14). Disque Q : - Q/Division/Doc_Entr/NI : 1,68 Go (5'755 fichiers, 24 dossiers). - Q/Division/C/SCO : 370 Ko (27 fichiers, 8 dossiers, non-accessible).		P

		- Q/Division/T/DAO : 9,59 Go (19'628 fichiers, 1534 dossiers). Q/Division/DRI :Bâtiments IB : non-accessible, vide ?		
		6. Définition des références relatives aux objets d'information à délivrer GED eDocs : Référence numérique existante avec #[n° id] Isilog : pas de n° de référence. Disque Q : pas de n° de référence.		P et A
		7. Choix des outils Outils du côté des archives : - Création du SIP via Docuteam Packer - Versement dans Docuteam cosmos.		P et A
		8. Description des objets d'information en référence à un dictionnaire de données et à un modèle Application de la liste ad-hoc des métadonnées pour l'archivage électronique des tpg.		P et/ou A
	<i>Création d'un dictionnaire</i>	9. Définition des classes d'objet et de leurs attributs		P et A
		10. Encodage du dictionnaire de données		P et A
	<i>Construction d'un modèle formel</i>	11. Définition du modèle de données	- dépend de la manière dont sont délivrées les données	P et A
		12. Elaboration d'un modèle de représentation		P et A
	<i>Formalisation des aspects légaux et contractuels</i>	13. Description des aspects légaux et contractuels		P et A
	<i>Définition des conditions de transfert des documents et des données</i>	14. Définition des modalités de versement des documents et des données (procédures de communication) Médium : transfert des documents à archiver dans un dossier pourvu se trouvant dans un disque réseau.	- différents scénarios : médium physique (disque dur externe, clé USB, cédérom...), réseau informatique - authenticité : somme de contrôles	P et A
		15. Définition du paquet d'information (Information qu'on veut verser pour archiver)		P et A
		16. Définition des sessions de transfert (fréquence : annuel, semestriel...) Annuel.	- fréquence (quotidien, hebdomadaire, mensuel...)	P et A
		17. Définition du test de transfert initial		P et A
		18. Identification des outils à utiliser durant la phase de transfert (plateforme de pérennisation et d'archivage, ISM) - Création du SIP via Docuteam Packer. - Versement dans Docuteam cosmos.		P et A
		19. Description des procédures de transfert (authenticité, fiabilité, traçabilité, intégrité) - Le producteur ou SIT extrait les documents. - Le producteur ou SIT met les documents sur un dossier d'un disque réseau pourvu pour le transfert.		P et/ou A

		- L'archiviste récupère les documents crée le SIP et verse les documents dans Docuteam cosmos.		
	Définition de la validation	20. Définition d'un plan de validation immédiat	- éléments : complétude (tous les objets ont été correctement versés), intégrité, conformité par rapport au modèle de données	A
		21. Définition d'un plan de validation reporté		A
		22. Définition des procédures de validation / non-validation des objets		P et A
		23. Définition du test de validation initial		P et A
		24. Identification des outils de validation		A
		25. Description des procédures de validation		A
	Planification du transfert	26. Définition d'une planification de référence		P et A
		27. Définition des procédures de non-respect de la planification		P et A
	Gestion des changements après la conclusion du protocole de versement		- raisons : changement d'infrastructure, nouveaux objets à verser, modification des ressources, changement légal... - procédure : identification de l'origine et des causes de la demande de changement, élaboration de différents scénarios pour gérer ce changement, évaluation de la faisabilité, du travail à réaliser et des coûts de chaque scénario, décision, définition et exécution du plan d'action pour implanter le changement	P et A
	Evaluation de la faisabilité, des coûts et des risques			P et A
PROTOCOLE DE VERSEMENT			- objets à verser (contenu des SIP, modèle de SIP, communauté(s)-cible(s), aspects légaux et contractuels) - définition des modalités de versement - définition des modalités de validation - gestion du changement - planification des versements	P et/ou A

Annexe 8 : Protocole de versement – eDocs

Auteur [Prénom Nom, fonction]
Date 12 juillet 2022
Version [N° de version]

Le présent document définit le protocole entre les domaines d'activité (DA) et les Archives des tpg pour le versement des documents en version finale de la GED eDocs sur la plateforme de pérennisation et d'archivage des Archives des tpg.

1. Contexte

La GED eDocs est l'outil de gestion électronique des documents, aussi utilisé comme un outil collaboratif aux tpg. Elle est destinée à contenir les versions abouties et finales des documents émanant de tous les domaines d'activité des tpg.

À l'heure actuelle, la GED contient plus de 307'000 fichiers avec une augmentation annuelle estimée à 20'000 fichiers¹⁶. L'ancienneté de eDocs au tpg laisse présager la possibilité de retrouver des documents remontant jusqu'aux années 1990.

Parmi cette masse documentaire, 181 documents et séries de documents listés en annexe sont à archiver.

Vu l'ancienneté supposée et le volume d'accroissement des documents sur cette application, il devient prioritaire de préparer la conservation de ce contenu.

Le protocole prévoit l'archivage de ces documents sous forme électronique, par extraction depuis la GED eDocs. Les documents sont versés dans la plateforme de pérennisation et d'archivage docuteam cosmos.

2. Information sur le producteur

Les documents sont produits par tous les domaines d'activité (DA) des tpg.

Transports publics genevois
Route de La-Chapelle 1
1212 Lancy

Personnes de contact :

- Interviewé 23, référent métier et Spécialiste en gestion documentaire.
- Nouveau-elle référent-e SIT.
- Répondants documentaires dans les DA.

3. Description des documents versés

3.1. Typologie des documents versés

Les documents versés sur la plateforme de pérennisation et d'archivage sont, entre autres, des PV de séances, des rapports (rapports d'activité, de sécurité, d'audit, de contrôles,

¹⁶ Entretien avec l'interviewé 22, Responsable Infrastructures et Opérations, Genève, 11 avril 2022.

de financement ou techniques), des correspondances, des politiques, des chartes, des règlements et manuels. La liste complète des documents à archiver est disponible en annexe. Ces documents sont aux formats : PDF, DOC(X), PPT(X) et XLS(X).

3.2. Période de création

La période de création des documents versés sur la plateforme de pérennisation et d'archivage peut varier de la date d'introduction de eDocs au sein des tpg dans les années 1990 à aujourd'hui.

3.3. Volume des données

Le volume exact des documents est à déterminer. Il faut communiquer tous les identifiants des fichiers afin de les retrouver dans la GED et lancer l'estimation du volume.

3.4. Accès

L'accès aux documents déposés sur la plateforme de pérennisation et d'archivage est en principe réservé à la consultation interne.

L'accès après leur sort final, est public à l'échéance des délais de protection (cf. extraits de la LArch en fin de document).

4. Versement

Le versement d'archives électroniques doit répondre à un certain nombre d'exigences archivistiques et techniques.

4.1. Exigences techniques

Les fichiers versés doivent être aux formats suivants : PDF/A ou XLSX.

Le nommage des fichiers est systématisé de la manière suivante :

- Numéro d'identifiant - aaaammjj - nom du document

4.2. Exigences archivistiques

Les exigences archivistiques comprennent la liste des métadonnées retenues suivante :

- Nom du producteur ;
- Rubrique du plan de classement ;
- Type de document ;
- Sous-type de document ;
- Intitulé du document ;
- Date de création du document ;
- Numéro unique.

Le cas échéant les métadonnées exigées sont au minimum le titre et la date.

4.3. Modalités de versement

Le processus prévu est le suivant :

- Extraction annuelle depuis la GED eDocs vers le dossier « transfert » du disque réseau Q ;
- Migration des documents en formats pérennes ;
- Création du paquet d'information (SIP) avec Docuteam Packer ;
- Versement des documents dans la plateforme de pérennisation et d'archivage.

5. Rôles et responsabilités

Les Archives des tpg

- Font part de la liste des documents souhaités pour l'archivage à la référente métier eDocs et aux répondants documentaires ;
- Créent le paquet d'information (SIP) ;
- Réalisent les opérations de versement de ces données sur la plateforme de pérennisation et d'archivage avec les répondants SIT ;
- Assurent l'utilisation, la sécurité et l'organisation des documents dans la plateforme de pérennisation et d'archivage.

Les répondants documentaires et la référente métier eDocs

- Répertorient les documents sollicités par les Archives et y communiquent les identifiants.

Les répondants SIT

- Effectuent l'extraction des documents sollicités à l'aide des identifiants ;
- Effectuent les migrations de format nécessaires ;
- Réalisent les opérations de versement de ces données sur la plateforme de pérennisation et d'archivage avec les archivistes ;
- Assurent l'entretien et le fonctionnement général de la plateforme de pérennisation et d'archivage.

6. Révision

Le présent protocole de versement peut être révisé en tout temps, notamment lorsqu'il s'agit de verser sur la plateforme de pérennisation et d'archivage de nouveaux types de documents.

Approbation des parties	
Date :	Date :
<p>[Signature]</p> <p>[Prénom et nom] Parties prenantes</p>	<p>[Signature]</p> <p>[Prénom et nom] Archiviste des tpg</p>

Loi sur les archives publiques (LArch) du 1^{er} décembre 2001, art. 12 :

« ³ Le délai général de protection est de 25 années à compter de la clôture du dossier. Le dernier apport organique est déterminant pour définir l'année au cours de laquelle les dossiers ont été clos.

⁴ Les documents classés selon des noms de personnes et qui contiennent des données personnelles sensibles ou des profils de la personnalité ne peuvent être consultés que 10 ans après le décès de la personne concernée, à moins que celle-ci n'en ait autorisé la consultation. Si la date de la mort est inconnue ou n'est déterminable que moyennant un travail disproportionné, le délai de protection expire 100 ans après la naissance. Si ni la date du décès, ni celle de la naissance ne peuvent être déterminées, le délai de protection expire 100 ans à compter de l'ouverture du dossier. »

Annexe 9 : Protocole de versement – Isilog

Auteur [Prénom Nom, fonction]
Date 12 juillet 2022
Version [N° de version]

Le présent document définit le protocole entre la Direction Générale (DG), le Conseil Administratif (CA) et les Archives des tpg pour le versement des documents d'Isilog sur la plateforme de pérennisation et d'archivage des Archives des tpg.

1. Contexte

Isilog est l'outil de suivi des fiches d'amélioration (FA) d'entreprise. Les non-conformités et les recommandations sont entrées dans le système pour être suivies¹⁷.

Au total, l'outil contiendrait une centaine de GB¹⁸, la quantité exacte de la volumétrie pour les données concernées par l'archivage reste à déterminer.

Isilog existe aux tpg depuis 2008, il n'a pas subi de mise à jour et ne peut être utilisé qu'avec l'ancienne version du navigateur Microsoft Internet Explorer. Par conséquent, celui-ci sera décommissionné à une date non-communicée. Les données à archiver pourraient remonter jusqu'à 2008.

Vu les caractéristiques de cet outil il est prioritaire de préparer la conservation de son contenu.

Le protocole prévoit l'archivage de ces documents sous forme électronique, par extraction depuis Isilog. Les documents sont versés dans la plateforme de pérennisation et d'archivage docuteam cosmos.

2. Information sur le producteur

Les producteurs des documents sont la Direction Générale (DG) et le Conseil Administratif (CA) des tpg.

Transports publics genevois
Route de La-Chapelle 1
1212 Lancy

Personnes de contact :

- Interviewé 1, référent métier.
- Interviewé 14, référent SIT.

3. Description des documents versés

3.1. Typologie des documents versés

Les documents versés sur la plateforme de pérennisation et d'archivage sont pour la Direction générale : le manuel de qualité, les fiches d'amélioration et de suivi des audits, les rapports d'audit des autorités de surveillance, les Rapport d'événement suite à un accident

¹⁷ Entretien avec l'interviewé 1, Responsable Management d'entreprise, Genève 18 mai 2022.

¹⁸ Entretien avec l'interviewé 14, Chef de projet IT, Genève, 25 mai 2022.

professionnel et la carte stratégique. Pour le Conseil administratif, il s'agit des rapports d'audit sur le fonctionnement financier et organisationnel.

Les audits peuvent être extraits avec des requêtes SQL sous forme de tables et les documents en pièces-jointes de ces audits sont aux formats : PDF, DOC(X), PPT(X), XLS(X).

3.2. Période de création

La période de création des documents versés sur la plateforme de pérennisation et d'archivage peut varier de la date d'introduction de Isilog au sein des tpg en 2008 à aujourd'hui.

3.3. Volume des données

Le volume exact des données à archiver reste à déterminer.

3.4. Accès

L'accès aux documents déposés sur la plateforme de pérennisation et d'archivage peut être limitée dans un premier temps, car Isilog contient certaines données sensibles et confidentielles.

L'accès après leur sort final est public à l'échéance des délais de protection (cf. extraits de la LArch en fin de document).

4. Versement

Le versement d'archives électroniques doit répondre à un certain nombre d'exigences archivistiques et techniques.

4.1. Exigences techniques

Les fichiers versés doivent être aux formats suivants : PDF/A ou XLSX pour les documents joints aux audits et en SIARD pour les données extraites avec les requêtes SQL.

Le nommage des fichiers est systématisé de la manière suivante :

- Numéro d'identifiant - aaaammjj - nom du document

4.2. Exigences archivistiques

Les exigences archivistiques comprennent la liste des métadonnées retenues suivante :

- Nom du producteur ;
- Rubrique du plan de classement ;
- Type de document ;
- Sous-type de document ;
- Intitulé du document ;
- Date de création du document ;
- Numéro unique.

Le cas échéant les métadonnées exigées sont au minimum le titre et la date.

4.3. Modalités de versement

Le versement à partir d'Isilog se fait en une fois.

Isilog est en phase de décommissionnement, une fois que le nouvel outil ou la nouvelle version prendra la relève, le processus prévu sera le suivant :

- Extraction annuelle vers le dossier « transfert » du disque réseau Q ;
- Migration des documents en formats pérennes ;
- Création du paquet d'information (SIP) avec docuteam Packer.
- Versement des documents dans la plateforme de pérennisation et d'archivage.

5. Rôles et responsabilités

Les Archives des tpg

- Font part de la liste des documents souhaités pour l'archivage aux référents métier et SIT ;
- Créent le paquet d'information (SIP) ;
- Réalisent les opérations de versement de ces données sur la plateforme de pérennisation et d'archivage avec les répondants SIT ;
- Assurent l'utilisation, la sécurité et l'organisation des documents dans la plateforme de pérennisation et d'archivage.

Les répondants SIT

- Effectuent l'extraction des documents sollicités ;
- Effectuent les migrations de format nécessaires ;
- Réalisent les opérations de versement de ces données sur la plateforme de pérennisation et d'archivage avec les archivistes ;
- Assurent l'entretien et le fonctionnement général de la plateforme de pérennisation et d'archivage.

6. Révision

Le présent protocole de versement peut être révisé en tout temps, notamment lorsqu'il s'agit de verser sur la plateforme de pérennisation et d'archivage de nouveaux types de documents.

Approbation des parties	
Date :	Date :
<div>[Signature]</div> <div>[Prénom et nom] Parties prenantes</div>	<div>[Signature]</div> <div>[Prénom et nom] Archiviste des tpg</div>

Loi sur les archives publiques (LArch) du 1^{er} décembre 2001, art. 12 :

« ³ Le délai général de protection est de 25 années à compter de la clôture du dossier. Le dernier apport organique est déterminant pour définir l'année au cours de laquelle les dossiers ont été clos.

⁴ Les documents classés selon des noms de personnes et qui contiennent des données personnelles sensibles ou des profils de la personnalité ne peuvent être consultés que 10 ans après le décès de la personne concernée, à moins que celle-ci n'en ait autorisé la consultation. Si la date de la mort est inconnue ou n'est déterminable que moyennant un travail disproportionné, le délai de protection expire 100 ans après la naissance. Si ni la date du décès, ni celle de la naissance ne peuvent être déterminées, le délai de protection expire 100 ans à compter de l'ouverture du dossier. »

Annexe 10 : Protocole de versement – Disque réseau Q

Auteur [Prénom Nom, fonction]
Date 12 juillet 2022
Version [N° de version]

Le présent document définit le protocole pour le versement des documents concernés par l'archivage se trouvant dans le disque Q vers la plateforme de pérennisation et d'archivage des Archives des tpg.

1. Contexte

Le disque réseau Q est à sa base prévu pour les dossiers temporaires et contient un dossier « transfert ». Il contient cependant aussi une multitude d'autres dossier dont les éléments à archiver suivants :

- Les notes d'information (NI) ;
- Les séances OPS (Opérations) hebdomadaires et le Service client opérationnel ;
- Les plans des véhicules ;
- Le dossier bâtiment avec les autorisations de construire.

À l'heure actuelle, les dossiers contenant les documents à archiver dans le disque réseau Q cumulent plus de 25'410 fichiers et 11 GB de données. Les documents remontent dans certains cas au début des années 2000.

Le protocole prévoit l'archivage de ces documents sous forme électronique, par récupération depuis le disque réseau. Les documents sont versés dans la plateforme de pérennisation et d'archivage docuteam cosmos.

2. Information sur le producteur

Les producteurs de ces documents sont les secteurs Exploitation (E) pour les notes d'informations, Technique (T) pour les plans de véhicules, Marketing Vente et Communication (MVC) pour les séances OPS et l'ancien domaine d'activité Infrastructure et Bâtiments (I&B) pour les autorisations de construire.

Transports publics genevois
Route de La-Chapelle 1
1212 Lancy

Les personnes de contact sont à déterminer.

3. Description des documents versés

3.1. Typologie des documents versés

Les dossiers et documents mentionnés plus haut incluent les formats suivants : DOC, DOCX, XLS, XLSX, PPT, RTF, TXT, PDF, JPEG, TIFF, HTML, MSG, DWG, DXF et IDW.

3.2. Période de création

Certains documents datent du début des années 2000 et s'étendent jusqu'à aujourd'hui. Les fichiers dans les Notes d'informations (NI) sont créés de 2001 à aujourd'hui et sont organisés par années, la liste des NI remonte jusqu'à 1994 mais les fichiers sont plus récents.

3.3. Volume des données

Le volume total des dossiers accessibles est de 11 GB.

3.4. Accès

L'accès aux dossiers et documents archivés sur la plateforme de pérennisation et d'archivage, après leur sort final, est public à l'échéance des délais de protection (cf. extraits de la LArch en fin de document).

4. Versement

Le versement d'archives électroniques doit répondre à un certain nombre d'exigences archivistiques et techniques.

4.1. Exigences techniques

Les fichiers versés doivent être aux formats recommandés pour la conservation en annexe. Les formats mentionnés plus haut (cf. 3.1. Typologie des documents versés) doivent être convertis selon les cas en PDF/A, TXT, XLSX, DWG, DXF, IDW, JPEG, TIFF, MSG et HTML.

Le nommage des fichiers est systématisé de la manière suivante :

- Numéro d'identifiant - aaaammjj - nom du document

4.2. Exigences archivistiques

Les exigences archivistiques comprennent la liste des métadonnées retenues suivante :

- Nom du producteur ;
- Rubrique du plan de classement ;
- Type de document ;
- Sous-type de document ;
- Intitulé du document ;
- Date de création du document ;
- Numéro unique.

Le cas échéant les métadonnées exigées sont au minimum le titre et la date.

4.3. Modalités de versement

Le processus prévu est le suivant :

- Récupération depuis le disque réseau Q ;
- Migration des documents en formats pérennes ;
- Création du paquet d'information (SIP) avec Docuteam Packer.
- Versement des documents dans la plateforme de pérennisation et d'archivage.

5. Rôles et responsabilités

Les Archives des tpg

- Récupèrent les fichiers disponibles sur le disque réseau Q ;
- Créent le paquet d'information (SIP) ;
- Assurent l'utilisation, la sécurité et l'organisation des documents dans la plateforme de pérennisation et d'archivage.

Les Archives avec l'aide d'un référent SIT

- Effectuent les migrations de format nécessaires ;
- Réalisent les opérations de versement de ces données sur la plateforme de pérennisation et d'archivage ;
- Assurent l'entretien et le fonctionnement général de la plateforme de pérennisation et d'archivage.

6. Révision

Le présent protocole de versement peut être révisé en tout temps, notamment lorsqu'il s'agit de verser sur la plateforme de pérennisation et d'archivage de nouveaux types de documents.

Approbation des parties	
Date :	Date :
<div>[Signature]</div> <div>[Prénom et nom] Parties prenantes</div>	<div>[Signature]</div> <div>[Prénom et nom] Archiviste des tpg</div>

Bases légales (extraits)

Loi sur les archives publiques (LArch) du 1^{er} décembre 2001, art. 12 :

« ³ Le délai général de protection est de 25 années à compter de la clôture du dossier. Le dernier apport organique est déterminant pour définir l'année au cours de laquelle les dossiers ont été clos.

⁴ Les documents classés selon des noms de personnes et qui contiennent des données personnelles sensibles ou des profils de la personnalité ne peuvent être consultés que 10 ans après le décès de la personne concernée, à moins que celle-ci n'en ait autorisé la consultation. Si la date de la mort est inconnue ou n'est déterminable que moyennant un travail disproportionné, le délai de protection expire 100 ans après la naissance. Si ni la date du décès, ni celle de la naissance ne peuvent être déterminées, le délai de protection expire 100 ans à compter de l'ouverture du dossier. »