



PROJET NUMERISATION DES DOCUMENTS CONSERVES A L’IST

**CCTP
ACCORD CADRE
PRESTATION DE NUMERISATION**

	Rédacteur	Vérificateurs		Approbateur
Fonction Nom	Caroline Martin	Responsable AQ étude		Responsable de l'étude
Visa				

HISTORIQUE

Version Révision	Date de mise en application	Cause et/ou nature de l'évolution
1.0	09/05/2025	Création
1.1	12/06/2025	Modification
1.2	19/08/2025	Validation
1.3	24/08/2025	Modifications et ajouts

SOMMAIRE

1 INTRODUCTION 4

2 DESCRIPTION DE LA PRESTATION ATTENDUE 4

2.1 PERSONNES RESSOURCES 4

2.2 OBJET DU MARCHE ET DESCRIPTION DU BESOIN 4

2.3 ENLEVEMENT ET RESTITUTION DES LOTS 5

2.4 NUMERISATION DES DOCUMENTS 5

2.5 NUMERISATION 3D D'OBJETS 7

2.6 NUMERISATION DE FILMS : 7

2.7 RECAPITULATIF 8

2.8 FOURNITURE DES METADONNEES 8

3 LIVRABLES 10

4 CRITERES DE VALIDATION 10

1 INTRODUCTION

L'ONERA souhaite sous-traiter la numérisation et l'océrisation de documents (publications scientifiques et rapports techniques) ainsi que de la numérisation d'objets en 3D et de films dans la perspective de leur communication sur son portail documentaire. Le titulaire ne conserve aucun droit sur le produit de la numérisation : fichiers sources, fichiers PDF, fichiers XML et métadonnées.

2 DESCRIPTION DE LA PRESTATION ATTENDUE

2.1 PERSONNES RESSOURCES

Des personnes référentes sont désignées au sein de la DIST (Direction de l'Information Scientifique et Technique de l'ONERA) pour assurer le suivi technique de l'exécution des prestations, selon les centres concernés : :

- Châtillon & Palaiseau
- Lille & Modane
- Toulouse

Leurs coordonnées professionnelles (nom, fonction et adresse électronique) seront transmises au titulaire lors de la notification de l'accord-cadre.

Le titulaire doit fournir un interlocuteur unique pour chaque bon de commande.

Si ce référent technique ne peut être le référent administratif, le titulaire devra également fournir le contact d'un référent administratif.

2.2 FORME, OBJET DE L'ACCORD-CADRE ET DESCRIPTION DU BESOIN

Le marché est un accord-cadre à bons de commande.

Il porte sur la prestation de numérisation des fonds scientifique documentaire et les rapports techniques pour les centres de Châtillon, Palaiseau, Modane, Toulouse et Lille, ainsi que d'objets type maquettes et des films.

Le nombre moyen de pages par documents est donné à titre informatif et n'est pas contractuel.

.

Le présent accord cadre est alloti comme suit :

- Lot 1 : Document type A4 A3-A5 simple (feuille à feuille) + document spécifique (photographies, microformes, diapositives, plans, calques, plaques de verre...),
- Lot 2 : Maquettes numérisation 3D,
- Lot 3 : Numérisation de documents **classifiés** (type A4 A3-A5 simple (feuille à feuille) et documents spécifiques (plans, calques, plaques de verre...),
- Lot 4 : numérisation de films (bobines + cassettes vidéo magnétiques).

2.3 ENLEVEMENT ET RESTITUTION DES LOTS

2.3.1 Pour les documents/objets sans mention de protection : Pour les lots 1, 2 et 4, les lots de documents/objets à numériser sont à retirer par La Poste, un service de messagerie, un transporteur, un coursier...sur le site ONERA de localisation des documents

- Châtillon à la Direction de l'Information Scientifique et technique (DIST) de l'ONERA
- Palaiseau à la Direction de l'Information Scientifique et technique (DIST) de l'ONERA
- Lille au centre de documentation du centre ONERA
- Toulouse au centre de documentation du centre ONERA

Des cartons vides et adaptés au transport de documents/objets sont fournis par le prestataire à la Direction de l'Information Scientifique et Technique (DIST) en amont de l'enlèvement.

Les documents/objets sont conditionnés en carton et préparés par la Direction de l'Information Scientifique et Technique (DIST). Le titulaire a la charge du transport aller, du reconditionnement après numérisation et du retour.

Des emballages réutilisables, en matériaux recyclés et recyclables sont à privilégier (cf. Article XII de l'accord-cadre).

Les assurances transports et liées à la prestation de numérisation sont à la charge du prestataire.

2.3.2 Pour les documents nécessitant une prestation sur place (lot 3 documents sous mention de protection et classifiés)

Le prestataire (et les personnels qu'il affecte à la réalisation des prestations) doit être habilité (Défense -cf. accord-cadre) et organise la prestation de numérisation sur site :

- Châtillon
- Palaiseau
- Toulouse
- Lille
- Modane

Les caractéristiques des documents sous mention de protection et classifiés répondent aux caractéristiques des lots 1, 2 et 4. Ces caractéristiques sont décrites dans les paragraphes suivants du présent cahier de charges.

L'organisation de la prestation sur site fait l'objet d'un suivi « ad hoc » en fonction du site concerné.

2.4 NUMERISATION DES DOCUMENTS

Caractéristiques générales :

Les documents numérisés livrés sont dans une résolution standard de 300 dpi.

Exceptionnellement, l'ONERA peut demander une résolution supérieure (jusqu'à 1200 dpi). Cette demande sera inscrite dans le fichier de récolement fourni à chaque lot de documents à numériser.

En cas de documents mixtes, l'ONERA précise les spécificités de la numérisation dans le fichier de récolement.

- les formats requis sont :
 - le format PDF/A-3a basé sur Pdf 1.7. Ces formats sont décrits dans la norme ISO 19005-3 (PDF/A-3) et ISO 32000-1
 - le format Tiff groupe 4 pour les sources noir et blanc (ou équivalent), (norme ISO 12639 : 2004)
 - le format JPEG (ou équivalent) pour les sources en niveau de gris et en couleur, (Norme ISO 10918),

Les engagements du prestataire :

Toutes les pages blanches non paginées doivent être supprimées, sauf précision contraire.

Selon indication dans le tableau de récolement, le format original des documents doit être respecté (A5, A4, A3,...). Les documents sont de format majoritairement A4.

Le cadrage est mono-page sauf précision contraire.

Les documents sont numérisés couvertures comprises, en couleur, en 300 dpi.

Les pages doivent être redressées, au besoin, ou réorientées.

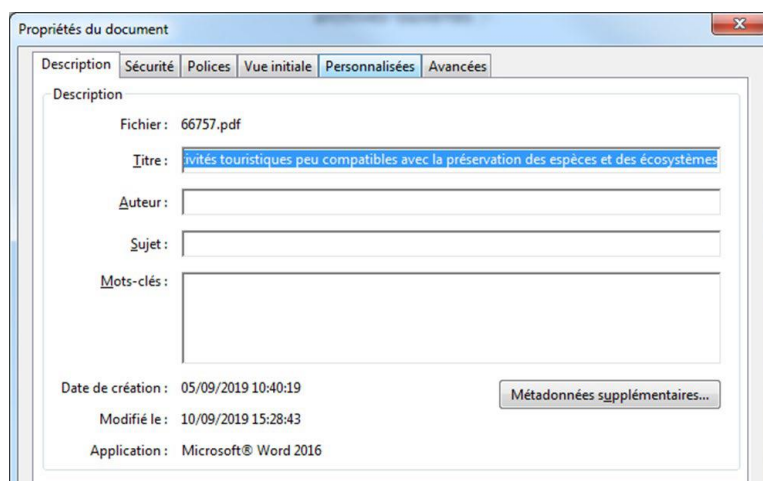
Les alignements doivent être redressés. La référence est la ligne de texte qui doit être horizontale, à la perpendiculaire des bords gauche-droit de la page.

Les bordures de page non nettes doivent être rognées.

Un traitement de reconnaissance de caractère (OCR) est appliqué à l'intégralité de chaque fichier. Pour le texte intégral, il est attendu une qualité proche du standard de l'édition, soit 1/10000 (1 erreur maximum sur 10 000 caractères). Le contrôle humain nécessaire à la qualité suppose une compétence linguistique (documents majoritairement en français et anglais). Il est attendu 100% d'exactitude dans les métadonnées (fichier XML), dans le texte des 1^{ère} et 4^{ème} de couverture et en page de garde du document numérisé (fichier PDF) ainsi que dans les propriétés du document PDF. Le titre du document est enregistré dans les propriétés du document PDF.

Le résultat est un fichier xml dont le contenu est inclus dans un fichier PDF multicouche pour l'affichage du texte en transparence avec le mode image et autoriser la recherche

Exemple :



Les engagements de l'ONERA :

Afin de permettre au prestataire d'établir le devis pour chaque lot de documents à numériser, l'ONERA s'engage à fournir pour chaque besoin (bon de commande) dans un fichier tableur (Excel ou équivalent) dit fichier de récolement, les éléments suivants :

- La sélection des documents à numériser, leur morphologie et leur présentation avec la mention « **A reconstruire** » en cas de massicotage et/ou la mention « **Ne pas massicoter** » le cas échéant;
- Pour chaque lot de documents à numériser, les métadonnées des documents concernés, afin qu'il soit procédé à leur injection dans les propriétés des documents.

Le fichier de récolement ainsi constitué sera le reflet de l'ordre de classement des documents physiques proposés à la numérisation.

Les fichiers sont nommés de manière incrémentale (le pas est de +1) en commençant par 001 pour le premier fichier. Ex :

- 001.pdf

- 002.pdf

- ...

2.5 NUMERISATION 3D D'OBJETS

Caractéristiques générales :

- Les objets seront numérisés en 3D selon le modèle 3D CAO avec reconstitution d'un objet.
- La précision métrique attendue > 0,5 mm
- La résolution du nuage de points attendu est 5 points/mm²
- Photogrammétrie intégrée ou texturation HD sur le maillage.
- Couleurs fidèles (balance des blancs, résolution > 8 Mpx).

Caractéristiques des fichiers à livrer :

Nuage de points : .e57, .xyz, .las → pour l'archivage scientifique.

Maillage 3D texturé : .obj + .mtl + textures en .jpg ou .png → pour la visualisation ou l'impression 3D.

Le format doit être compatible avec les logiciels utilisés par l'ONERA (notamment CATIA...). Les fichiers sont organisés et nommés clairement

La documentation jointe comprend :

- Rapport technique complet :
 - Méthodologie (type de scanner, photogrammétrie, conditions de lumière).
 - Date et lieu.
 - Résolution, précision, marges d'erreur.
- Métadonnées intégrées dans les fichiers ou en annexe : auteurs, copyright, objectifs, licence de diffusion (en lien avec les données du fichier de récolement communiqué par l'ONERA)

L'ONERA fournit au prestataire un fichier de récolement indiquant toutes les informations nécessaires à la réalisation de la prestation ainsi que les métadonnées à intégrer dans chaque fichier à livrer.

2.6 NUMERISATION DE FILMS :

- Films bobines (8MM, 16MM, 35MM)
- Cassettes vidéo (VHS, VHS-C)

Caractéristiques générales :

- Définition de 720p à 1080p selon le degré de qualité du film (full HD)

Le format doit être compatible avec les logiciels utilisés par l'ONERA (notamment Windows Media Player, VLC, QuickTime...). Les fichiers sont organisés et nommés clairement.

La documentation jointe comprend :

- Rapport technique complet :
 - Méthodologie (type de scanner, photogrammétrie, conditions de lumière).
 - Date et lieu.
 - Résolution, précision, marges d'erreur.
- Métadonnées intégrées dans les fichiers ou en annexe : auteurs, copyright, objectifs, licence de diffusion (en lien avec les données du fichier de récolement communiqué par l'ONERA)

L'ONERA fournit au prestataire un fichier de récolement indiquant toutes les informations nécessaires à la réalisation de la prestation ainsi que les métadonnées à intégrer dans chaque fichier à livrer

2.7 RECAPITULATIF

Type de documents	Format de numérisation attendu niveau de gris	Format de numérisation attendu couleur	Format de numérisation attendu noir & blanc
Ouvrages / rapports / articles papier (A4-A3)	PDF /A-3a (norme ISO 19005-3 " ISO 32000-1)	PDF /A-3a (norme ISO 19005-3/ ISO 32000-1)	PDF /A-3a (norme ISO 19005-3/ ISO 32000-1)
Photographies papier	JPEG (ou équivalent)	JPEG (ou équivalent)	TIFF groupe 4(ou équivalent)
Diapositives / phototypes (24x36, 6x6 et 6x9)	JPEG (ou équivalent)	JPEG (ou équivalent)	TIFF groupe 4 (ou équivalent)
Cartes / plans en calque + papier et plié + rouleau (A2 à A0)	JPEG (ou équivalent)	JPEG (ou équivalent)	TIFF groupe 4 (ou équivalent)
Objets / Maquettes en bois, balsa, acier, composite de différentes dimensions (petites dimensions à 2m de longueur sur 50 cm de largeur sur 1m de hauteur)	obj / ou équivalent lisible par un logiciel CAO / JPG ou PNG pour la visualisation (norme ISO 15948)	obj / ou équivalent lisible par un logiciel CAO / JPG ou PNG pour la visualisation (norme ISO 15948)	obj / ou équivalent lisible par un logiciel CAO / JPG ou PNG pour la visualisation (norme ISO 15948)
Films 8MM 16MM 35MM cassettes VHS, VHS-C	MP4 (MPEG-4) (compatible VLC, Windows Media Player et Quick Time / navigateurs Chrome/firefox/safari et You tube) Norme ISO 14496	MP4 (MPEG-4) (compatible VLC, Windows Media Player et Quick Time / navigateurs Chrome/firefox/safari et You tube) Norme ISO 14496	MP4 (MPEG-4) (compatible VLC, Windows Media Player et Quick Time / navigateurs Chrome/firefox/safari et You tube) Norme ISO 14496

2.8 FOURNITURE DES METADONNEES

L'ensemble des métadonnées listées ci-dessous est collecté dans le texte intégral et enregistré dans un fichier XML par lot de documents numérisés. Un exemple est fourni en annexe. Le format des champs est présenté ci-dessous :

- **fichier** : nom du fichier PDF correspondant
- **titre_fr** : version française du titre du document
- **titre_en** : version anglaise du titre du document
- **auteur** : nom de la / des personnes, prénom de la / des personnes, auteur moral à l'origine de la rédaction du document
- **classification** : Diffusion Restreinte, Spécial France, Confidentiel Industrie...
- **type_de_document** : Thèse/Mémoire/HDR/Rapport de stage/Rapports techniques
- **annee** : année de publication au format YYYY
- **langue_texte_integral** : fra ou eng. Langue de rédaction du texte du document numérisé
- **affiliations** : L'affiliation est la structure de recherche à laquelle appartient un auteur au moment où il a produit le document.
- **n° rapport** : n° du rapport

- **localisation** : Châtillon, Palaiseau, Toulouse, Lille, Modane
- **Mots-clefs** : mots-clefs
- **Nombre de pages** : Nombre de pages du document

Les champs titre, résumé et auteurs sont contrôlés manuellement : si l'OCR contient des erreurs, ces erreurs sont toutes corrigées dans les métadonnées fournies (fichier XML). Ces données doivent également être exactes dans le texte intégral (1^{ère} et 4^{ème} de couverture, page de garde). Sur ces métadonnées, il est attendu 100 % d'exactitude.

Les titres, s'ils sont en majuscules dans le document PDF, sont transformés en minuscules dans le fichier XML à l'exception des sigles, de la première lettre du titre, de la première lettre des noms propres et des acronymes.

Le fichier XML fourni est valide : il est possible de l'ouvrir et de le visualiser dans Google Chrome sans erreur. Dans cette perspective, les caractères interdits en XML sont encodés (ex : remplacement des < en <). Une attention particulière doit être apportée aux symboles (physique, mathématiques...)

Les fichiers doivent impérativement être corrigés avant livraison.

Les fichiers numérisés seront publiés sur le portail documentaire interne ONERA pour être visualisés en ligne. La taille des fichiers PDF livrés doit donc être optimisée à l'aide d'une option équivalente à la fonctionnalité « Réduire la taille du fichier » du logiciel Acrobat. La taille des fichiers PDF ne devra pas excéder 10 Mo par fichier de 100 pages.

Un seul fichier XML **contient l'ensemble des métadonnées d'un même lot de documents**. Le lien entre PDF et métadonnées correspondantes dans le fichier XML est établi par le champ « fichier » du fichier XML où le nom du fichier PDF est renseigné, ex <fichier>001.pdf</fichier>.

Pour les objets numérisés en 3D :

Les métadonnées suivantes sont intégrées dans les fichiers livrés:

- **Auteurs** : nom de la / des personnes, prénom de la / des personnes
- **année** de fabrication de l'objet
- **copyright**, ONERA
- **Résumé** contexte de l'objet,
- **localisation** : Châtillon, Palaiseau, Fauga-Mauzac, Toulouse, Lille, Modane
- **Mots-clefs** : mots-clefs
- **licence de diffusion**

Les métadonnées seront présentes dans le fichier de récolement transmis au prestataire.

Pour les films :

Les métadonnées suivantes sont intégrées dans les fichiers XML livrés :

- **Auteurs** : nom de la / des personnes, prénom de la / des personnes
- **Titre du film**
- **année** de réalisation
- **copyright**, ONERA
- **Résumé** contexte de réalisation,
- **localisation** : Châtillon, Palaiseau, Toulouse, Lille, Modane
- **Mots-clefs** : mots-clefs
- **licence de diffusion**

Les métadonnées seront présentes dans le fichier de récolement transmis au prestataire.

3 LIVRABLES

Le livrable est composé des éléments suivants :

1. Les originaux reconditionnés en carton dans l'ordre défini par le tableau de récolement,
2. Pour les documents : un fichier PDF/A 3a par document numérisé
3. Pour les documents : Le tableau de récolement complété avec les noms des fichiers PDF reportés par le titulaire dans la colonne « nom de fichier » de la ligne correspondant au document ainsi que d'éventuelles remarques
4. Dans le cas de la numérisation 3D, la restitution des objets
5. Dans le cas des prestations sur site (lot 3 uniquement), les documents sont repositionnés dans les rayonnages et boîtes d'origine.
6. Chaque document massicoté doit être individuellement conditionné en chemise élastiquée ou enveloppe.
7. Un unique fichier XML par lot de documents numérisés contenant les métadonnées les décrivant, livré au moyen d'un service WEB
8. Pour les objets numérisés en 3D, les fichiers ODJ et .IGES avec la description en format ASCII des objets numérisés en 3D et des fichiers JPG/PNG pour la visualisation ainsi que le rapport technique complet
9. Pour les films un fichier de recollement et les lots de fichiers MP4 transmis sur un serveur web
10. Une copie de sécurité doit être conservée par le titulaire tant que la livraison n'est pas effective et que la Direction de l'Information Scientifique et Technique (DIST) n'a pas effectué les opérations de vérifications et admis les prestations, cette copie pourra être exigée en cas de nécessité, elle sera détruite après admission.

4 TRANSFERT DES CENTRES DANS LE CADRE DU PROJET PRISME

Dans le cadre du projet PRISME, l'ONERA prévoit le regroupement de plusieurs de ses centres en Île-de-France, avec des fermetures progressives prévues à l'horizon 2027–2029.

A ce jour, il est d'ores et déjà acté que le centre de Châtillon fermera. Les activités actuellement localisées sur ce site seront transférées vers le centre de Palaiseau, qui deviendra le site de référence pour ces prestations.

Ainsi, si le titulaire est amené à récupérer des documents à Châtillon au cours de l'exécution de l'accord-cadre, et que le site ferme entre-temps, la récupération et la restitution de ces documents devront alors s'effectuer à Palaiseau, sans incidence sur le contenu technique des prestations attendues (le ou les bons de commande concernés préciseront ce point).

L'ONERA attire l'attention du titulaire sur le fait que d'autres évolutions du périmètre d'intervention (fermetures ou transferts d'activités) peuvent intervenir dans le cadre du projet PRISME. Ces changements sont susceptibles d'avoir un impact sur l'organisation des prestations, notamment en termes de localisation, sans pour autant modifier les exigences techniques prévues dans le présent cahier des charges.

5 EXIGENCES TECHNIQUES

Pour les lots 1 et 3 :

- Tous les documents/objets originaux sont restitués intègres, selon les critères indiqués dans le tableau de récolement.
- La Direction de l'Information Scientifique et Technique effectuera un premier contrôle qualité portant sur le nombre de documents restitués, le nombre de fichiers, leur nommage, leur poids, leur validité...

- Un second contrôle sera effectué par sondage : nombre, ordre, orientation, cadrage, netteté des pages, qualité de l'image, fidélité de l'océrisation, exactitude des métadonnées...
- Les délais convenus en accord avec le titulaire et figurant dans le bon de commande sont respectés. Ce délai sera fixé en fonction du volume à traiter. Le délai maximum de réalisation entre commande et livraison est de 4 mois.
- Un fichier PDF par document et un unique fichier XML par lot de documents (un même bon de commande) sont livrés pour chaque bon de commande.
- Les fichiers PDF et XML correspondent aux attentes décrites dans le tableau de récolement et respectent les spécifications du présent cahier des charges.

Pour le lot 2 :

- Un fichier au format ..obj pour ouverture par un logiciel CAO
- Un rapport de contrôle 3D, sous forme de PDF, présentant des cartographies 3D et des contrôles géométriques 3D

Pour le lot 4 :

- Un fichier format excel de recensement des films traités avec les métadonnées
- Lots de fichiers format MP4