Cahier des clauses techniques particulières (CCTP)

— CCTP LOT 03 —

DEPLOIEMENT, DEVELOPPEMENT ET MAINTENANCE D'UNE PLATEFORME NUMAHOP : LOGICIEL OPEN SOURCE DE GESTION DE PROJETS DE NUMERISATION

INFRASTRUCTURE NUMÉRIQUE DU PORTAIL « BIBLIOTHÈQUES VIRTUELLES HUMANISTES » (BVH)

EQUIPEX+ BIBLISSIMA+

« Fairisation des données BVH vers les portails et outils Biblissima et partenaires »

Biblissima+ bénéficie d'une aide de l'Etat gérée par l'ANR au titre du Programme d'investissements d'avenir intégré à France 2030, portant la référence ANR-21-ESRE-0005.

1. OBJET DU LOT 03 3

2. CONTEXTE TECHNIQUE ET FONCTIONNEL 4

2.1. ARCHITECTURE ACTUELLE ET ARCHITECTURE CIBLE 4
2.2. NUMAHOP ET LES FLUX DE NUMERISATION BVH 5
2.3. REFERENTIELS NUMAHOP A PRENDRE EN COMPTE 6

3. OBJECTIFS FONCTIONNELS DU LOT 03 6

3.1. DEPLOYER ET STABILISER LA PLATEFORME NUMAHOP SUR LA NOUVELLE INFRASTRUCTURE 7
3.2. ASSURER LA MAINTENANCE CORRECTIVE 7
3.3. [PSE OBLIGATOIRE] ASSURER LA MAINTENANCE EVOLUTIVE 7
3.4. [PSE OBLIGATOIRE] INTERFACER NUMAHOP AVEC LE STOCKAGE OBJET S3 (LOT 01) 8
3.5. [PSE OBLIGATOIRE] INTERFACER NUMAHOP AVEC LE SOCLE APPLICATIF (POSTGRESQL, API REST, IIIF — LOT 01) 8
3.6. ASSURER LA TRAÇABILITE ET LE SUIVI DES FLUX DE NUMERISATION 8
3.7. GARANTIR L'INTEROPERABILITE DES DONNEES BVH, LEUR REUTILISABILITE ET LA PERENNITE DES EXPORTS 9

4. PERIMETRE DES PRESTATIONS ATTENDUES 9

4.1. REUNION DE LANCEMENT ET PRISE DE CONNAISSANCE 9
4.2. AUDIT TECHNIQUE ET FONCTIONNEL INITIAL 9
4.3. DEPLOIEMENT, CONFIGURATION ET ENVIRONNEMENTS DE L'INSTANCE NUMAHOP CIBLE 10
4.4. PARAMETRAGE FONCTIONNEL DES FLUX BVH 10
4.5. MAINTENANCE CORRECTIVE ET PREVENTIVE 11
4.6. [PSE OBLIGATOIRE] MAINTENANCE EVOLUTIVE ET DEVELOPPEMENTS SPECIFIQUES 11
4.7. [PSE OBLIGATOIRE] INTERFAÇAGE AVEC LE STOCKAGE OBJET S3 12
4.8. [PSE OBLIGATOIRE] INTERFAÇAGE AVEC LE SOCLE APPLICATIF (POSTGRESQL, API REST, IIIF, OAI) 12
4.9. IMPORTS DES NOTICES ET FICHIERS, ET ARTICULATION AVEC LE LOT 01 13
4.10. GENERATION DES EXPORTS ET ARTICULATION AVEC LE LOT 01 13
4.11. CONTROLES, VALIDATION ET SUIVI QUALITE 14

5. EXIGENCES TECHNIQUES TRANSVERSALES AU LOT 03 14

6. LIVRABLES ATTENDUS DU LOT 03 15

7. MODALITES DE DEPLOIEMENT SPECIFIQUES AU LOT 03 16

8. TEST & RECETTES SPECIFIQUES AU LOT 03 16

9. SECURITE 17

10. DOCUMENTATION ET TRANSFERT DE COMPETENCES 17

1. OBJET DU LOT 03

Le présent lot a pour objet le déploiement, la maintenance corrective et évolutive, ainsi que l'adaptation fonctionnelle et technique de la plateforme NumaHOP, logiciel libre de gestion de chaînes de numérisation, au sein de l'infrastructure numérique BVH.

NumaHOP couvre l'ensemble de la chaîne de numérisation : import des notices bibliographiques, constitution de lots et de trains de numérisation, constats d'état, réception et contrôle des fichiers numérisés livrés par les prestataires, contrôle qualité, structuration des unités documentaires, génération des fichiers techniques (OCR, METS, ALTO, images dérivées) et export vers les diffuseurs et les entrepôts d'archivage pérenne.

Les BVH disposent déjà d'une instance dockerisée et modifiée de NumaHOP, assimilée à un fork ou à une adaptation locale, hébergée et maintenue sur les serveurs de l'Université de Tours au sein du DataCentre régional Centre-Val de Loire. Le titulaire devra auditer cette instance, identifier les écarts avec le dépôt GitHub de référence de NumaHOP, proposer une stratégie de mise à niveau, puis réaliser les opérations nécessaires pour disposer d'une instance maintenable, documentée et intégrée à l'infrastructure cible BVH.

L'offre porte prioritairement sur l'audit, le réalignement, la stabilisation, le paramétrage, la maintenance corrective, les tests, la documentation et le transfert de compétences. Les évolutions fonctionnelles propres aux besoins BVH, les connecteurs avancés pour l'interfaçage avec le stockage S3 et le socle applicatif du lot 01, ainsi que la maintenance évolutive, sont à décrire et à chiffrer séparément en option.

Le lot 03 s'appuie sur et alimente le socle applicatif du lot 01, sans se substituer à lui. Le lot 01 conserve la responsabilité du stockage objet S3, de l'API REST, de l'exposition IIIF et OAI-PMH, des manifests, de l'indexation, des identifiants publics, etc. Le lot 03 doit en revanche préparer, contrôler ou transmettre les ressources et métadonnées nécessaires aux traitements des numérisations.

Le lot 03 constitue une brique de production et de traitement. Il devra permettre :

- la stabilisation de l'instance NumaHOP sur la VM dédiée, administrée par l'université de Tours
- le suivi et la traçabilité des flux de numérisation (statuts, historique, indicateurs)
- la documentation, les tests, la recette et le transfert de compétences
- la maintenance corrective de la plateforme (anomalies, mises à jour de sécurité, compatibilité des dépendances)

Le lot comprend également les trois prestations supplémentaires optionnelles suivantes, qui seront chiffrées séparément par les candidats :

- [PSE obligatoire] l'interfaçage de NumaHOP avec le stockage objet S3 déployé dans le cadre du lot 01

- [PSE obligatoire] l'interfaçage de NumaHOP avec le socle applicatif du lot 01 : base PostgreSQL « Catafonds », API REST, serveur IIIF, pour la synchronisation des métadonnées, des identifiants et des états de numérisation
- [PSE obligatoire] la maintenance évolutive et les développements spécifiques nécessaires aux usages BVH (paramétrages métier, connecteurs, formats d'export)

2. CONTEXTE TECHNIQUE ET FONCTIONNEL

NumaHOP est un progiciel libre (SpringBoot/Java, build Maven) développé initialement par la bibliothèque Sainte-Geneviève, la BULAC et la bibliothèque de Sciences Po Paris, aujourd'hui utilisé par une quinzaine d'établissements documentaires. Il repose nativement sur une base de données MariaDB, un moteur d'indexation Elasticsearch dédié, et s'appuie sur des outils tiers pour le traitement des fichiers : ImageMagick et ExifTool pour le contrôle et la manipulation des images, Tesseract pour l'OCR. Il prend en charge l'import de notices aux formats UNIMARC, EAD, Dublin Core (simple ou qualifié), CSV et JSON, notamment depuis le Sudoc et Calames, ainsi que la génération de fichiers OCR, METS et d'images dérivées en sortie.

Le programme BVH met en place une infrastructure numérique modulaire destinée à gérer, traiter, publier, interroger et valoriser des documents patrimoniaux de la Renaissance. La chaîne de traitement doit permettre de prendre en compte les images de documents numérisés, les métadonnées descriptives et techniques, les fichiers dérivés, les fichiers OCR ou structures, notamment les sommaires et tables des matières.

Dans cette architecture, NumaHOP doit permettre l'import, le contrôle, le suivi, l'enrichissement fonctionnel et l'export des données et fichiers issus des flux de numérisation. Les sorties de NumaHOP ont vocation à être exploitées par le lot 01 pour le stockage, l'exposition des données, la diffusion IIIF, l'indexation et la préparation des services consommés par les lots 01 et 02.

L'objectif n'est pas de faire de NumaHOP le référentiel central des métadonnées BVH. La base PostgreSQL « Catafonds » et les services du lot 01 restent les composants centraux pour l'exposition, la publication et l'interopérabilité. NumaHOP doit être positionné comme outil de gestion de flux, de contrôle et de préparation des ressources issues de la numérisation.

2.1. Architecture actuelle et architecture cible

Une VM dédiée à l'outil NumaHOP a été déployée à la DSI de l'université de Tours au sein du DataCentre Centre-Val de Loire, sous Debian 12. Cette VM héberge une instance de test dockerisée et modifiée de NumaHOP (fork, avec corrections et adaptations locales réalisées en interne avec le soutien du CETU ILIAD3). Cette instance a permis à l'équipe BVH d'effectuer des premiers tests et évaluer les besoins de fonctionnalités et paramétrages spécifiques pour BVH. L'instance NumaHOP cible, à mettre à jour et à optimiser dans le cadre du présent lot, se substituera à l'instance test, sur cette même VM.

L'architecture cible repose sur une circulation maîtrisée des données entre les trois lots : le lot 03 gère les flux de numérisation et prépare les ressources ; le lot 01 centralise, stocke, expose et indexe les données ; le lot 02 consomme les services exposés pour l'affichage public, la recherche et la consultation IIIF. L'architecture cible devra donc permettre :

- le dépôt automatisé des images et fichiers dérivés validés dans le bucket S3 du lot 01, dans l'arborescence et selon le plan de nommage définis avec l'équipe BVH
- la transmission ou la synchronisation des métadonnées techniques et des états de numérisation vers la base PostgreSQL / l'API REST du lot 01
- la disponibilité, pour le serveur IIIF notamment, des images et fichiers dérivés produits par NumaHOP, en cohérence avec le mécanisme de génération des manifests du lot 01
- le versionnement, le suivi des anomalies et la documentation spécifique aux BVH sur le GitLab BVH

Les accès devront prioritairement passer par les services et API exposés par le socle applicatif BVH. Les accès directs de NumaHOP à PostgreSQL, au stockage objet S3 ou aux index de recherche (si PSE du lot 01 retenue) devront être justifiés, documentés et validés par le maître d'ouvrage.

2.2. NumaHOP et les flux de numérisation BVH

NumaHOP organise la production autour de projets, eux-mêmes subdivisés en lots puis en trains de numérisation, avec gestion des constats d'état (avant envoi, à réception chez le prestataire, après numérisation, au retour) et des échanges avec un ou plusieurs partenaires ou prestataires de numérisation externes.

Les données et flux concernés par le présent lot comprennent notamment :

- les notices bibliographiques importées dans NumaHOP : les métadonnées requises issues de la base PostgreSQL « Catafonds », qui constitue la base principale des métadonnées descriptives
- les constats d'état des documents physiques
- les images issues des campagnes de numérisation (fichiers maîtres et dérivés) avec les métadonnées encapsulées
- les sommaires/tables des matières préétablis (tabulaires ou arborescents) selon les normes de structuration BVH
- les fichiers techniques produits : OCR, METS, ALTO
- les statuts d'avancement des lots et trains de numérisation
- les exports vers les diffuseurs (bibliothèque numérique BVH, portails partenaires) et vers l'archivage pérenne (CINES ou solution équivalente)

Les flux prioritaires pour BVH sont :

- la reprise de numérisations anciennes, actuellement déjà publiées mais pas encore exposées sous IIIF
- le traitement des numérisations anciennes, pas encore traitées ni publiées
- le traitement de nouvelles numérisations

Le prestataire devra veiller à la cohérence d'ensemble des flux de données entre les fichiers importés dans NumaHOP et ceux exportés : les métadonnées descriptives et techniques, les identifiants pérennes, les formats des images, des manifests IIIF, etc.

Le lot 03 ne porte pas sur le nettoyage scientifique, l'arbitrage documentaire ou la restructuration conceptuelle des données BVH. Le titulaire devra toutefois signaler les incohérences et anomalies qui empêchent l'import, le contrôle, le traitement ou l'export correct des données dans la chaîne NumaHOP.

2.3. Référentiels NumaHOP à prendre en compte

À ce stade, la documentation utilisateur et technique de NumaHOP reste partielle. Le site public NumaHOP demeure une source de référence pour la description générale du logiciel, de ses modules et de ses usages, tandis qu'un travail collaboratif est en cours pour structurer la documentation dans le référentiel GitHub officiel.

Référence	**URL**	**Usage attendu par le prestataire**
Site public NumaHOP	https://www.numahop.fr	Documentation fonctionnelle et utilisateur de référence historique
Documentation publique	https://numahop.github.io/Documentation/	Manuels utilisateur et développeur, notes de version depuis la version 26.03
Code source de la documentation	https://github.com/NumaHOP/Documentation	Documentation technique en cours de structuration
Organisation GitHub NumaHOP	https://github.com/NumaHOP	Dépôt de référence pour le code, les *issues*, *pull requests* et *releases*
Release 26.03	https://github.com/NumaHOP/NumaHOP/releases/tag/v26.03	Point de référence récent pour le code et les notes de version

Les candidats devront distinguer ce qui est documenté, ce qui est disponible dans le code, ce qui sera actif dans l'instance BVH et ce qui relèvera d'une évolution annoncée ou non stabilisée. En cas de divergence entre documentation, code et comportement de l'instance, le constat effectif sera documenté par le titulaire dans le rapport d'audit.

3. OBJECTIFS FONCTIONNELS DU LOT 03

Sur la base de l'existant, et en interaction étroite avec les prestations réalisées dans le cadre du lot 01, le lot 03 rendra opérationnelle la chaine de traitement des numérisations en tenant compte des spécificités et besoins du programme BVH pour le traitement, la publication et à terme l'archivage des documents patrimoniaux.

Le lot 03 doit aussi permettre à l'équipe BVH de reprendre la maîtrise technique de l'instance NumaHOP et de ses procédures de mise à jour.

Le lot 03 devra répondre aux objectifs fonctionnels suivants :

3.1. Déployer et stabiliser la plateforme NumaHOP sur la nouvelle infrastructure

Le titulaire réalisera un audit technique et fonctionnel de l'instance NumaHOP actuellement expérimentée par les BVH. Cet audit identifiera l'état de l'installation, les paramétrages existants, les dépendances, les risques techniques et les conditions de mise à niveau.

Le titulaire proposera et mettra en œuvre une stratégie de mise à niveau permettant de réaligner l'instance BVH avec le code désormais disponible sur le dépôt Git de référence de NumaHOP. Cette opération devra préserver les configurations et paramétrages utiles aux BVH (peu nombreux à ce stade et n'incluant pas de données), ou prévoir leur reprise à l'issue d'une migration.

Le titulaire livrera une instance stabilisée, documentée, reproductible et maintenable : versionnement, documentation, scripts de déploiement, variables d'environnement, sauvegarde, restauration, tests et transfert de compétences (*cf.* CCTP communs aux lots 01 à 03).

3.2. Assurer la maintenance corrective

Le titulaire assurera la correction des anomalies identifiées, l'application des correctifs de sécurité et le maintien en conditions opérationnelles de la plateforme, dans le respect des versions communautaires de NumaHOP et de ses dépendances (Java/SpringBoot, MariaDB, Elasticsearch, ImageMagick, ExifTool, Tesseract, etc.).

3.3. [PSE obligatoire] Assurer la maintenance évolutive

Le prestataire fera évoluer NumaHOP pour répondre aux besoins métier spécifiques des BVH : paramétrage des workflows, des profils utilisateurs, des modèles de constats d'état, des formats d'export, ainsi que le développement de connecteurs ou d'adaptations non couverts nativement par le progiciel communautaire.

Les développements spécifiques seront proposés en amont à la communauté NumaHOP par le biais de son association. Le cas échéant, les développements seront remontés au dépôt communautaire GitHub ; les adaptations locales non reversées seront mises à disposition sur le dépôt GitLab BVH dédié à NumaHOP.

Les fonctionnalités supplémentaires envisagées (à développer ou déjà en cours de développement selon la feuille de route du logiciel tenue par l'Association NumaHOP) dans le cadre de cette consultation, pourront par exemple porter sur : l'import et l'export de manifests IIIF, la gestion et la saisie des sommaires/tables des matières arborescents directement dans un visualiseur IIIF intégré à NumaHOP, l'annotation, etc.

3.4. [PSE obligatoire] Interfacer NumaHOP avec le stockage objet S3 (lot 01)

Le titulaire développera ou adaptera les connecteurs permettant à NumaHOP de déposer les images sources, les fichiers dérivés et les fichiers techniques (OCR, METS, ALTO) dans les volumes du stockage objet S3 (MinIO ou solution assimilée) dédiés, en lecture pour les fichiers sources à traiter et en écriture pour les fichiers produits après traitement, en cohérence avec l'organisation des buckets définie dans le lot 01.

3.5. [PSE obligatoire] Interfacer NumaHOP avec le socle applicatif (PostgreSQL, API REST, IIIF — lot 01)

Le titulaire mettra en œuvre les mécanismes de synchronisation entre NumaHOP et le socle applicatif du lot 01 : transmission des métadonnées descriptives et techniques, des statuts de numérisation, des identifiants internes et, le cas échéant, des identifiants ARK, afin de permettre la génération des manifests IIIF et la publication des notices sur le portail (lot 02).

Le mécanisme technique (API REST synchrone, webhook, export/import périodique, table de correspondance partagée...) sera arbitré en concertation avec le titulaire du lot 01 et le maitre d'ouvrage en comité technique (COTECH) et comité de pilotage (COPIL). Il sera documenté conjointement, afin d'éviter toute divergence d'interprétation.

3.6. Assurer la traçabilité et le suivi des flux de numérisation

Le titulaire garantira la traçabilité des différentes étapes de la chaîne de numérisation (import, constat d'état, envoi, réception, contrôle, validation, export) et la disponibilité de ces informations pour le suivi de projet par l'équipe BVH.

Le titulaire vérifiera et, si nécessaire, paramétrera les modules NumaHOP utiles aux flux BVH : imports, métadonnées, fichiers images, contrôles, exports, structuration des lots, production de fichiers dérivés et articulation avec le stockage et les services du lot 01.

Le titulaire accompagnera le maitre d'ouvrage et sera force de proposition dans l'établissement d'une feuille de route avec priorisation et évaluation des développements supplémentaires futures (hors prestation), utiles pour renforcer et automatiser les contrôles qualité à l'import (nommage, séquençage, intégrité des formats, résolution, complétude des lots) au-delà des contrôles natifs de NumaHOP.

En outre, un interfaçage entre NumaHOP et la plateforme open-source Arkindex (Teklia), utilisée pour le traitement automatisé des images patrimoniales par IA (segmentation, reconnaissance d'écriture, indexation) est souhaité à terme par le maître d'ouvrage, afin d'articuler les flux de numérisation avec les traitements IA disponibles dans l'écosystème BVH.

3.7. Garantir l'interopérabilité des données BVH, leur réutilisabilité et la pérennité des exports

Dans la limite des prestations retenues, le titulaire contribuera par les solutions et développements mis en œuvre, à l'interopérabilité et à la FAIRisation des données BVH, c'est-à-dire à leur capacité à être trouvées, consultées, reliées, échangées et réutilisées par les services du portail et de l'infrastructure BVH, par ses partenaires ou par des services tiers.

Cette exigence ne constitue pas une prestation autonome distincte, mais un principe transversal applicable à l'ensemble des développements.

Le titulaire s'assurera en outre que les exports produits par NumaHOP (METS, ALTO, Dublin Core, images dérivées) restent conformes aux standards attendus et compatibles avec les besoins de diffusion et d'exposition (portail BVH, IIIF, OAI-PMH) et d'archivage pérenne (CINES ou solution équivalente retenue par la maîtrise d'ouvrage).

4. PÉRIMÈTRE DES PRESTATIONS ATTENDUES

4.1. Réunion de lancement et prise de connaissance

Le titulaire organisera une réunion de lancement propre à ce lot avec le maître d'ouvrage et les interlocuteurs techniques concernés. Cette réunion permettra de préciser les accès, les environnements, le calendrier, les priorités, les risques identifiés, les documents disponibles et les modalités de suivi.

Le titulaire produira un compte rendu de lancement, un planning prévisionnel et une liste des informations manquantes. Cette liste distinguera les informations bloquantes pour l'audit, celles nécessaires à la mise à niveau, et celles qui pourront être complétées pendant le paramétrage. Elle devra aussi identifier les dépendances avec les lots 01 et 02, notamment pour les flux de données, les espaces de stockage, les API et la recette.

Le titulaire aura accès aux dépôts Git, modèle de données, tests réalisés par la maitrise d'ouvrage et à toute documentation existante et utile à ce cadrage initial.

Une méthodologie agile sera privilégiée.

4.2. Audit technique et fonctionnel initial

Le titulaire réalisera un audit court de l'existant technique afin d'identifier :

- les écarts avec la version communautaire de référence
- l'état de la VM dédiée et des dépendances installées (MariaDB, Elasticsearch, ImageMagick, ExifTool, Tesseract, JVM)
- les paramétrages métier existants (projets, workflows, profils)
- les connecteurs ou scripts déjà développés localement, le cas échéant

- les risques techniques et les prérequis de déploiement

A l'issue de l'audit, le titulaire proposera une trajectoire de mise à niveau. Cette trajectoire précisera la version cible, la branche, le tag, la *release* ou le *commit* retenu, les prérequis techniques, les migrations nécessaires, les impacts sur les données, les impacts sur les paramètres, les évolutions de dépendances et les mesures de sécurisation avant intervention. Après validation par le maître d'ouvrage, il réalisera la mise à niveau de l'instance et documentera les opérations effectuées.

Aucune montée de version ne sera réalisée sur la seule base d'une branche de développement non stabilisée, sauf validation explicite du maître d'ouvrage. Le titulaire privilégiera une version identifiée, traçable et documentée.

4.3. Déploiement, configuration et environnements de l'instance NumaHOP cible

Le titulaire assurera le déploiement de l'instance cible sur la VM dédiée, dans un environnement conteneurisé Docker conformément aux exigences du CCTP commun. Il justifiera et documentera toute contrainte technique de NumaHOP rendant la conteneurisation partielle ou impossible pour certains composants.

Le titulaire précisera les éléments qui relèvent du dépôt GitHub officiel NumaHOP et ceux qui relèvent du dépôt GitLab BVH. Les développements locaux, configurations propres aux BVH et les procédures d'exploitation seront livrés dans l'environnement de versionnement BVH, tandis que les contributions génériques éventuelles à NumaHOP suivront les pratiques de contribution du projet open source.

Le titulaire établira et livrera des procédures reproductibles : installation, configuration, mise à jour, redémarrage, sauvegarde, restauration, contrôle après déploiement et retour arrière. Les variables d'environnement seront séparées du code source. Les secrets, jetons, mots de passe ou clés techniques ne devront pas figurer dans les dépôts de code.

L'ouverture de ports, les sauvegardes système, le monitoring et l'obtention des certificats relèvent de la DSI de l'Université de Tours.

4.4. Paramétrage fonctionnel des flux BVH

Le titulaire vérifiera, adaptera et documentera le paramétrage de NumaHOP pour les flux BVH. Les choix retenus seront formalisés dans un cahier de paramétrage validé par le maître d'ouvrage.

Ce paramétrage tiendra compte des deux familles de besoin prioritaires : la reprise ou le traitement de numérisations préexistantes, et la gestion de nouvelles campagnes de numérisation. Les choix retenus seront formalisés dans un cahier de paramétrage validé par le maître d'ouvrage.

Le paramétrage portera sur la structuration des projets, lots et unités documentaires, les rôles utilisateurs, les statuts de workflow, les imports de métadonnées et de fichiers, les contrôles qualité,

les exports, les règles de nommage, les emplacements de dépôt et les traitements associés aux images ou aux fichiers dérivés.

Le titulaire proposera une correspondance explicite entre les concepts NumaHOP et les concepts de modèle de données BVH. Cette correspondance ne devra pas contredire le modèle de données « Catafonds », mais permettre de suivre de manière opérationnelle les flux de numérisation.

4.5. Maintenance corrective et préventive

Le titulaire assurera, pendant la durée du marché (jusqu'à la date limite fixée par les pièces administratives de la consultation) et à compter de la recette définitive, la correction des anomalies bloquantes et non bloquantes selon des délais de traitement à définir, ainsi que le suivi des versions communautaires de NumaHOP et l'application des correctifs de sécurité.

Cette maintenance sera décrite comme un service d'assistance et d'intervention encadrée. Les candidats préciseront dans leur offre les niveaux de sévérité des anomalies et les délais de résolution proposés (GTI/GTR), les plages horaires de couverture de la maintenance, les modalités de sollicitation (ticket GitLab, astreinte...).

Le périmètre de la maintenance couvrira : l'application, les conteneurs, les dépendances, la base, les connecteurs.

Niveau	**Définition indicative**	**Traitement attendu**
Critique	Instance indisponible ou blocage complet d'un flux prioritaire	Prise en charge rapide, correction prioritaire
Majeure	Fonction essentielle dégradée, import/export ou contrôle bloqué partiellement	Analyse, correction ou proposition de solution de contournement
Mineure	Anomalie sans blocage majeur ou défaut documentaire	Correction programmée dans le cadre de la maintenance
Évolution	Nouveau besoin ou changement fonctionnel	Chiffrage séparé, hors maintenance corrective sauf accord

4.6. [PSE obligatoire] Maintenance évolutive et développements spécifiques

En prestation supplémentaire à chiffrer, les candidats proposeront les modalités d'une maintenance évolutive relative aux développements spécifiques qu'ils auront identifiés comme nécessaires à l'adaptation de NumaHOP, pour répondre aux besoins BVH. Ces développements spécifiques porteront sur les connecteurs avec le lot 01, les exports vers le stockage objet S3 ou services API, la prise en charge de structures documentaires propres aux BVH, ou la gestion de formats encore insuffisamment stabilisés dans l'instance cible, etc.

Chaque développement fera l'objet d'une description fonctionnelle, d'une analyse d'impact, d'une estimation de charge, d'un chiffrage séparé, de tests, d'une documentation et d'une stratégie de maintenance, dans le respect de l'architecture communautaire du logiciel, et en évitant toute divergence non maintenable avec le tronc commun (fork non documenté).

Les développements envisageables concerneront :

- des connecteurs ou scripts d'échange avec le lot 01
- des exports vers le stockage objet S3, IIIF, API ou autres composants du socle applicatif du lot 01
- l'adaptation de flux de métadonnées et formats d'import de notices (CSV, DublinCore, JSON) aux sources BVH (« Catafonds »)
- l'adaptation de formats d'export (METS, ALTO, dérivés) aux besoins d'exposition, de publication et d'archivage
- des fonctions relatives aux sommaires, tables des matières, manifests IIIF ou API NumaHOP

Les développements génériques seront proposés à la communauté NumaHOP sous forme de contribution open-source. Les fonctionnalités annoncées, en cours de développement ou insuffisamment stabilisées dans NumaHOP ne seront pas considérées comme acquises. Le titulaire vérifiera leur disponibilité dans la version cible et, le cas échéant, assurera leur intégration dans le cadre de la maintenance soit corrective (offre de base), soit évolutive (option de prestation supplémentaire).

4.7. [PSE obligatoire] Interfaçage avec le stockage objet S3

Le titulaire développera et intégrera les connecteurs nécessaires pour que NumaHOP :

- accède en lecture aux images sources à traiter, le cas échéant déposées dans un volume de stockage S3 dédié
- dépose en écriture, après traitement et validation dans NumaHOP, les images et fichiers dérivés dans le volume de stockage S3 de la collection concernée, compatible avec les services et protocoles d'exposition (IIIF notamment)
- respecte le plan de nommage et l'organisation des buckets définis dans le cadre du lot 01

4.8. [PSE obligatoire] Interfaçage avec le socle applicatif (PostgreSQL, API REST, IIIF, OAI)

Le titulaire développera et intégrera les connecteurs nécessaires pour mettre en œuvre :

- la transmission et la synchronisation des métadonnées techniques produites par NumaHOP (états de numérisation, identifiants, informations de constat d'état) vers la base PostgreSQL ou l'API REST du lot 01
- la réception, le cas échéant, des métadonnées descriptives ou des identifiants (ARK ou solution de redirection/gestion d'URL retenu) nécessaires à l'alimentation des exports NumaHOP
- la coordination avec le titulaire du lot 01 et le maitre d'ouvrage pour la définition du contrat d'interface (format d'échange, fréquence, mécanisme technique retenu)

4.9. Imports des notices et fichiers, et articulation avec le lot 01

Le titulaire identifiera, vérifiera ou adaptera les modalités et formats d'import des données dans NumaHOP et des fichiers nécessaires à l'alimentation des flux BVH, depuis les sources retenues par le maître d'ouvrage. Ces éléments comprennent notamment les fichiers tabulaires d'import de recollement ou de suivi, les notices et métadonnées descriptives, les images sources ou fichiers maîtres, les sommaires ou tables des matières, les fichiers dérivés, les fichiers techniques associés éventuels, ainsi que les rapports de livraison utiles au contrôle qualité, etc.

Les imports de notices ou métadonnées devront être examinés en tenant compte des formats pris en charge ou documentés par NumaHOP (CSV, MARC XML, UNIMARC, EAD, JSON et mappings vers Dublin Core ou Dublin Core qualifié). Pour BVH, ces formats ne doivent pas être considérés comme tous obligatoires.

En concertation avec le maître d'ouvrage, et en tenant compte des PSE retenues ou non dans les lots 01 et 03, le titulaire vérifiera les imports pertinents, disponibles et utiles à l'instance cible.

Lorsque les données proviennent du lot 01, l'import devra privilégier les services exposés par le socle applicatif : métadonnées descriptives ou techniques, images sources, fichiers maîtres et dérivés, informations nécessaires à l'articulation avec « Catafonds » ou le lot 01. Tout export direct depuis PostgreSQL ou tout dépôt direct dans le stockage objet S3 sera justifié, documenté et validé conjointement entre les parties afin d'éviter des duplications de données ou des contournements du socle applicatif.

4.10. Génération des exports et articulation avec le lot 01

De même que les imports, les exports issus de NumaHOP s'insèreront dans le flux de données global BVH sans duplication inutile de fichier et sans contournement du socle applicatif.

Le titulaire garantira la production, par l'instance NumaHOP BVH, des fichiers d'export nécessaires à l'exposition et à la publication sur le portail BVH (images et fichiers de données et métadonnées dérivés) et, le cas échéant, à l'archivage pérenne, en assurant leur conformité aux formats documentés et attendus par les services consommateurs (lot 01, archivage).

Les exports pourront concerner les images ou fichiers produits, les fichiers maîtres ou dérivés, les fichiers OCR, METS/ALTO, les métadonnées descriptives et techniques, les rapports de traitement, les sommaires ou tables des matières, les informations nécessaires à la génération ou à la vérification des manifests IIIF, etc.

L'export des données sera donc privilégié vers les services disponibles du socle applicatif. Les sorties seront utilisables par le stockage S3, les services API, IIIF, OAI-PMH du socle applicatif et le cas échéant, les traitements associés aux OCR ou aux structures documentaires (sommaires/tables des matières).

Le titulaire sera attentif à ne pas confondre la capacité éventuelle de NumaHOP à produire ou gérer des éléments IIIF avec le périmètre du socle applicatif du lot 01, qui demeure responsable de l'exposition IIIF et des manifests dans l'architecture BVH. Toute fonctionnalité IIIF dans NumaHOP devra être présentée comme un flux amont ou un complément, et non comme un remplacement du service IIIF préexistant dans le lot 01.

Les modalités d'archivage des données BVH ne sont pas encore validées à ce jour par le maitre d'ouvrage. Le dépôt pour l'archivage pérenne au CINES ou dans Nakala est à l'étude.

4.11. Contrôles, validation et suivi qualité

Le titulaire vérifiera et paramétrera les fonctions de contrôle disponibles dans NumaHOP pour les flux BVH. Il distinguera les contrôles automatisables, les contrôles assistés par l'outil et les contrôles manuels relevant de l'équipe BVH.

Le périmètre des contrôles restera réaliste et ne vise pas à développer un système de contrôle qualité spécifique dans le cadre de ce marché.

Les contrôles attendus porteront sur la présence et validation des fichiers, du nommage, des formats, des caractéristiques techniques des images, les métadonnées techniques, ainsi que sur la cohérence entre métadonnées et fichiers, la complétude des livraisons et les erreurs d'import ou de traitement.

Les formats d'images à contrôler prioritairement sont : TIFF, JPEG, JPEG2000.

Les contrôles qui nécessitent un développement spécifique devront être signalés et, le cas échéant, justifiés et chiffrés par les candidats.

5. EXIGENCES TECHNIQUES TRANSVERSALES AU LOT 03

Les composants livrés dans le cadre du lot 3 respecteront les exigences techniques communes décrites dans le CCTP commun.

Le prestataire respectera plus spécifiquement pour ce lot 03, les exigences suivantes :

- conteneurisation Docker de l'instance NumaHOP et de ses dépendances (MariaDB, Elasticsearch, outils de traitement d'image), dans la mesure des contraintes techniques du progiciel
- documentation complète des écarts éventuels entre l'instance BVH et la version communautaire de NumaHOP
- respect des formats d'échange standards (METS, ALTO, Dublin Core, UNIMARC, JSON)
- sécurisation des échanges avec le stockage objet S3 et le socle applicatif du lot 01
- compatibilité avec les contraintes des serveurs de l'établissement

- livraison du code et de la documentation spécifique dans le dépôt GitLab de l'équipe BVH
- maintien de la cohérence entre les données NumaHOP, les données du socle applicatif et les ressources publiées via IIIF

6. LIVRABLES ATTENDUS DU LOT 03

Les livrables seront déposés et versionnés en priorité dans le dépôt GitLab de l'équipe BVH, dans des formats éditables, documentés et exploitables par l'équipe projet. Lorsque cela est pertinent, les livrables seront également accompagnés d'exemples, de fichiers de configuration, de scripts, de procédures d'exécution et de tests.

Les livrables associés aux prestations supplémentaires éventuelles (PSE) seront clairement identifiés dans l'offre du candidat et feront l'objet d'un chiffrage distinct.

Les livrables attendus porteront notamment sur :

N° livrable	Prestation concernée	Livrables attendus
L3-01	Lancement	Compte-rendu de lancement et planning prévisionnel
L3-01	Audit technique initial	Note d'audit de l'existant technique NumaHOP (version, dépendances, paramétrages, écarts avec le tronc communautaire)
L3-02	Audit technique initial	Note de cadrage technique et planning prévisionnel validés par le maître d'ouvrage
L3-03	Déploiement	Dockerfiles, fichiers de configuration, docker-compose ou dispositif équivalent pour NumaHOP et ses dépendances
L3-04	Déploiement	Procédure de migration ou de bascule vers la nouvelle instance, incluant les mesures de continuité de service
L3-05	Maintenance corrective	Registre des anomalies traitées et documentation des correctifs appliqués
L3-06	Maintenance évolutive, si PSE retenue	Code source des évolutions et développements spécifiques réalisés
L3-07	Maintenance évolutive, si PSE retenue	Documentation des évolutions, paramétrages métier et écarts avec la version communautaire
L3-08	Interfaçage S3/MinIO, si PSE retenue	Connecteurs de dépôt/lecture des fichiers dans le stockage objet, documentés
L3-09	Interfaçage socle applicatif, si PSE retenue	Connecteurs ou mécanismes de synchronisation avec PostgreSQL / l'API REST du lot 01, documentés
L3-10	Interfaçage socle applicatif, si PSE retenue	Documentation du contrat d'interface entre NumaHOP et le lot 01 (formats, fréquence, mécanisme technique)
L3-11	Imports	Paramétrage ou adaptation des imports (formats et fichiers) : notices, modèles, etc.
L3-12	Exports	Exemples d'exports valides (METS, ALTO, OCR, images dérivées) sur corpus pilote

L3-13	Versionnement, GitLab	Organisation du dépôt, branches, tags, README et procédures de livraison
L3-14	Sécurité	Documentation des paramètres de sécurité, comptes techniques et droits d'accès
L3-15	Tests et recette	Protocoles de test et rapport de recette
L3-16	Sauvegarde et restauration	Procédures de sauvegarde et de restauration des composants NumaHOP (base, index, fichiers)
L3-17	Documentation générale	Documentation technique complète du lot 3 : architecture, installation, exploitation, maintenance
L3-18	Rapport final	Rapport final de livraison synthétisant les prestations réalisées, choix techniques et limites

7. MODALITÉS DE DÉPLOIEMENT SPÉCIFIQUES AU LOT 03

Les modalités générales de déploiement, de mise en recette, de mise en production, de sauvegarde préalable et de retour arrière sont définies dans le CCTP commun.

Pour le lot 3, le titulaire assurera le déploiement sur la VM dédiée à NumaHOP en coordination avec l'équipe BVH sur la VM dédiée à NumaHOP, et en articulation avec la VM du socle applicatif (lot 01) pour les échanges de fichiers et de métadonnées.

Le titulaire devra documenter les prérequis, les dépendances avec le lot 01, les étapes de déploiement, les sauvegardes préalables, les points de contrôle, les procédures de vérification après déploiement et les modalités de retour arrière.

Toute interruption de service après la mise en production de la nouvelle instance NumaHOP, sera planifiée et validée en amont avec le maître d'ouvrage.

8. TEST & RECETTES SPÉCIFIQUES AU LOT 03

La recette du lot 3 devra vérifier le bon fonctionnement de l'ensemble des composants livrés, sur la base d'un corpus pilote fourni par la maîtrise d'ouvrage.

Les tests devront notamment porter sur :

- le déploiement Docker de l'instance NumaHOP
- le dépôt et la lecture de fichiers dans le stockage S3/MinIO
- la synchronisation des métadonnées et des statuts avec le socle applicatif du lot 01

- la production des exports (METS, ALTO, OCR, dérivés) et leur conformité aux formats attendus
- la sauvegarde et la restauration
- la documentation fournie

9. SÉCURITÉ

Les exigences générales de sécurité sont définies dans le CCTP commun. Le titulaire les appliquera à l'ensemble des composants relevant du lot 3, et notamment :

- sécuriser les accès à l'interface d'administration de NumaHOP et aux comptes des partenaires et prestataires de numérisation externes
- sécuriser les échanges de fichiers et de métadonnées avec le stockage S3 et le socle applicatif du lot 01
- gérer de manière sécurisée les variables sensibles, mots de passe et jetons techniques
- documenter les droits d'accès différenciés entre équipe BVH, prestataires de numérisation et services techniques

10. DOCUMENTATION ET TRANSFERT DE COMPÉTENCES

Le titulaire fournira une documentation complète, en français ou en anglais, exploitable par l'équipe BVH, et prévoira une session de transfert de compétences pour les administrateurs et utilisateurs référents BVH.

La documentation devra permettre à l'équipe BVH de comprendre l'installation, de suivre les procédures courantes, d'identifier les points de vigilance et de dialoguer efficacement avec la DSI (ou un prestataire de maintenance ultérieur). La documentation couvrira l'architecture de l'installation, les prérequis, les procédures d'installation, de mise à jour, de sauvegarde et de restauration, le paramétrage fonctionnel, les imports, contrôles et exports, les anomalies connues, les limites identifiées et les recommandations de maintenance.

Une session de transfert de compétences pourra être proposée par les candidats à destination des administrateurs et utilisateurs référents.