ARNiA
Agence régionale du numérique et de l'intelligence artificielle

Opérateur public de solutions numériques

Cahier des clauses particulières

ARNia 1050S – Fiabilisation de la géolocalisation des établissements de Bourgogne-Franche-Comté

IDéO BFC : référentiel régional des entreprises et activités

Table des matières

1. Introduction 2
   1.1 Présentation de l'ARNia 2
   1.2 Contexte de l'achat 2
   1.3 Objectifs de la prestation 3
   1.4 Dimensionnement 3
2. Prestations attendues 3
   2.1 Présentation du processus de géolocalisation 3
   2.2 Installation et environnement 5
   2.3 Intégration des données 6
   2.4 Qualification et sélection 6
   2.5 Stockage, historique et sélection 7
   2.6 Tests et réception 8
   2.7 Transfert de compétences 8
3. Calendrier souhaité 8
4. Livrables 8
5. Modalités de réponse et de sélection 9
   5.1 Echanges électroniques 9
   5.2 Proposition financière 9
   5.3 Proposition technique 10
   5.4 Critères de sélection 10
6. Clauses administratives et financières 10

1. INTRODUCTION

1.1 Présentation de l'ARNia

Forte de plus de **1 800 adhérents**, l'Agence Régionale du Numérique et de l'Intelligence Artificielle (ARNia) est un acteur incontournable de la transition numérique auprès des collectivités territoriales, et tout particulièrement du bloc communal de la Région Bourgogne-Franche-Comté.

DICO

GROUPEMENT D'INTÉRÊT PUBLIC (GIP)
Un GIP exerce ses missions dans le cadre de son objet d'intérêt général. A caractère non lucratif, il agit en faveur des membres du groupement.

Ce groupement d'intérêt public (GIP) apporte des solutions quotidiennes aux évolutions des usages numériques. Elle fournit à ses adhérents, et singulièrement aux collectivités qui ne disposent d'aucune expertise numérique, des services simples, efficaces et au moindre coût (cybersécurité, dématérialisation, respect et de protection des données, etc.). Elle accompagne également les réflexions nationales et le déploiement des évolutions règlementaires ou de solutions innovantes (numérique responsable, IA, territoires intelligents).

En plus de la plateforme de logiciels pour ses adhérents, l'ARNia opère également 3 missions d'intérêt général **ouvertes à l'ensemble de l'écosystème local** :

- IDéO BFC, le dispositif de partage des données et de la connaissance en Bourgogne-Franche-Comté (ideo.arnia-bfc.fr)
- MedNum BFC : la mission pour la médiation numérique en Bourgogne-Franche-Comté (www.mednum-bfc.fr)
- CSIRT-BFC : le centre régional de cybersécurité (https://www.csirt-bfc.fr/)

1.2 Contexte de l'achat

Dans le cadre d'IDéO BFC, l'ARNia pilote la construction d'un référentiel régional partagé des entreprises et activités, à partir de SIRENE et de données partenaires. Trois chantiers sont engagés : géolocalisation, effectifs et catégorisation. La présente consultation porte uniquement sur la géolocalisation.

Les localisations issues du fichier Insee « Géolocalisation des établissements du répertoire Sirene pour les études statistiques » présentent des niveaux de précision variables. Des partenaires régionaux disposent de positions corrigées, parfois issues de relevés de terrain. L'enjeu est de confronter ces données, conserver la meilleure localisation disponible et pouvoir réexécuter le traitement lors des mises à jour.

1.3 Objectifs de la prestation

La présente consultation a donc pour objet :

- Déployer une chaîne de traitement reproductible sur l'infrastructure ARNia.
- Importer et rapprocher les localisations par SIRET.
- Appliquer des règles partagées de qualité et de sélection de la localisation de référence.
- Conserver les sources, les corrections et la traçabilité des choix.
- Automatiser une exécution mensuelle et permettre sa reprise par l'ARNia.

1.4 Dimensionnement

Le marché porte sur une prestation forfaitaire comprenant l'installation, le développement, les tests, la mise en production et le transfert de compétences.

Le montant maximum de la prestation est fixé à 12 000 € HT.

Le marché prendra fin à l'issue de la réception des prestations, après validation des livrables par l'ARNia.

Une période de garantie de bon fonctionnement de 6 mois est prévue à compter de la réception définitive. Durant cette période, le prestataire devra corriger, sans coût supplémentaire, toute anomalie liée aux développements ou aux configurations réalisés dans le cadre du marché. Aucune prestation de maintenance récurrente n'est prévue.

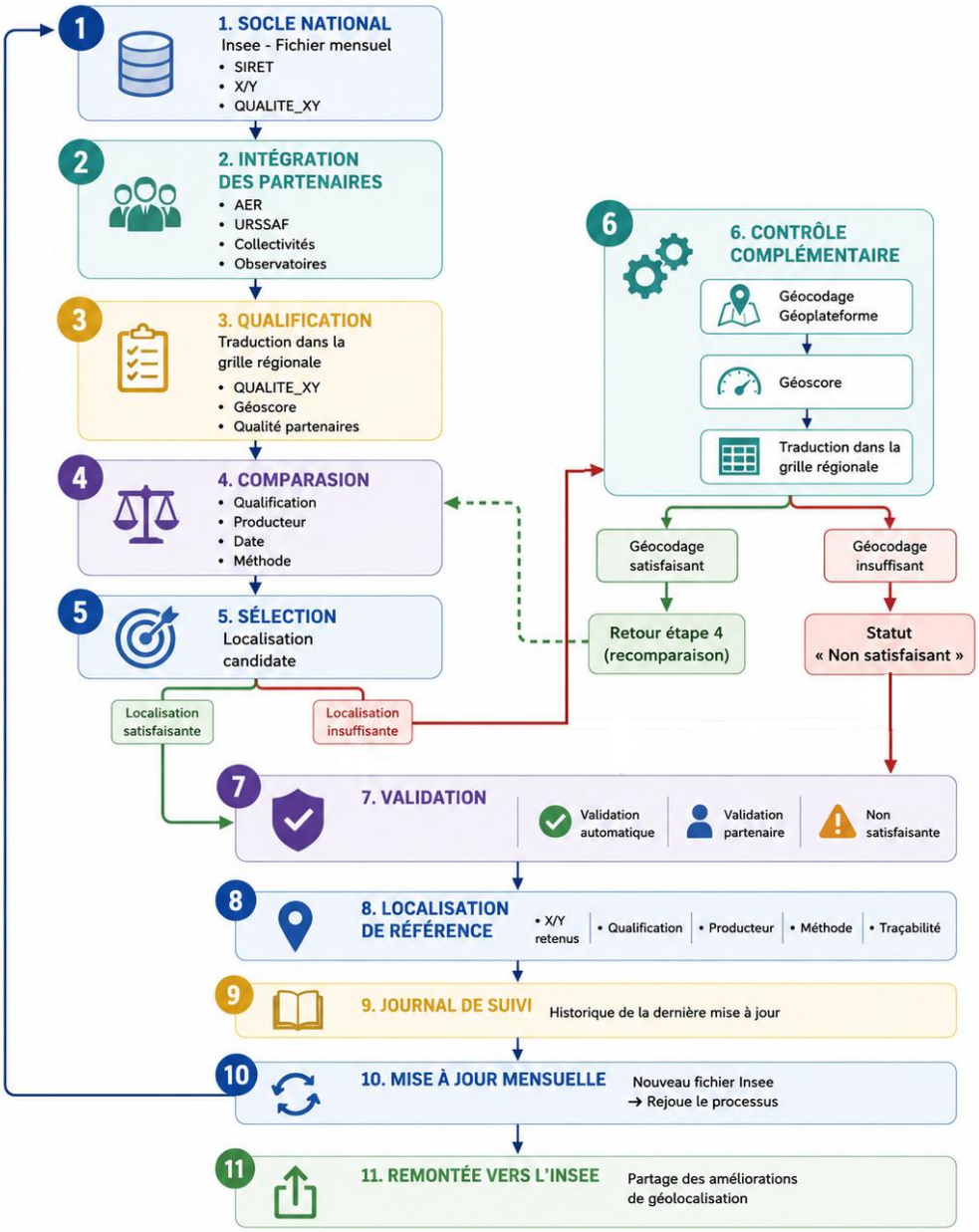
Périmètre : établissements de Bourgogne-Franche-Comté. La solution doit traiter le stock régional complet et ses mises à jour mensuelles. Le prestataire précisera les prérequis techniques et son estimation de charge. Aucun développement d'interface de consultation publique n'est demandé.

2. PRESTATIONS ATTENDUES

2.1 Présentation du processus de géolocalisation

Le processus de géolocalisation repose sur la confrontation des données de référence de l'Insee avec différentes sources nationales et régionales, afin de retenir la localisation la plus fiable pour chaque établissement.

Le schéma ci-dessous présente les principales étapes du processus envisagé.



Ce schéma présente les principes fonctionnels attendus. Le prestataire reste libre de proposer l'architecture technique et les outils nécessaires à leur mise en œuvre.

La qualification des localisations repose sur une grille régionale commune, définie par IDéO BFC et présentée en annexe 1.

Cette grille distingue deux dimensions complémentaires :

- **La précision géographique (P0 à P5)** : elle caractérise le niveau de positionnement de l'établissement, depuis l'absence de localisation jusqu'au bâtiment ou local effectivement occupé.

- **La fiabilité (F0 à F3)** : elle caractérise le niveau de confiance accordé à la localisation, en fonction de sa cohérence et des contrôles ou validations réalisés.

Ces deux dimensions sont indépendantes et doivent être conservées pour chaque localisation candidate.

Le prestataire devra intégrer cette grille dans la chaîne de traitement, assurer la correspondance avec les indicateurs de qualité des différentes sources et permettre l'évolution des règles de qualification sans modification majeure de la solution.

2.2 Installation et environnement

Le prestataire installe et configure PostgreSQL/PostGIS ainsi que les composants nécessaires à l'exécution et à l'automatisation des traitements sur l'infrastructure mise à disposition par l'ARNia. Il propose et justifie les outils retenus, en privilégiant une solution simple à exploiter et à maintenir.

L'ARNia souhaite privilégier l'utilisation de logiciels libres et open source pour l'ensemble des composants de la solution.

Le prestataire devra favoriser des outils pérennes, documentés et disposant d'une communauté active, afin de garantir l'autonomie de l'ARNia, la maintenabilité et l'évolutivité de la chaîne de traitement.

Le recours à des solutions propriétaires reste possible, sous réserve d'une justification technique et de l'absence de coûts de licence récurrents non prévus dans le marché.

L'ARNia mettra à disposition une machine virtuelle sous Ubuntu 26 ou Windows Server 2022, selon les besoins du prestataire et après validation par l'équipe infrastructure.

Dans sa réponse, le prestataire devra préciser les prérequis techniques suivants :

- Système d'exploitation : choix entre les deux environnements proposés.
- CPU, RAM et stockage : ressources nécessaires au fonctionnement de la solution et au traitement du stock régional complet.
- Ports et flux réseau : ports utilisés par les éventuelles interfaces web et d'administration, ainsi que les flux nécessaires au fonctionnement de la solution.
- DNS : éventuel besoin d'un sous-domaine pour accéder aux interfaces web.
- Comptes et droits : droits nécessaires à l'installation et au fonctionnement des composants applicatifs.

L'installation sera réalisée à l'aide d'un compte temporaire *tmp_install*, disposant des droits d'administration nécessaires (sudo sous Linux). Ce compte sera révoqué par

l'ARNia à l'issue de l'installation. Le prestataire devra garantir le fonctionnement de la solution avec un compte applicatif dédié, sans droits d'administration permanents.

Les accès distants seront assurés via le bastion Guacamole de l'ARNia, avec des comptes nominatifs fournis aux intervenants du prestataire.

Les prérequis techniques et les modalités d'accès seront validés par l'équipe infrastructure de l'ARNia avant le démarrage des prestations.

2.3 Intégration des données

La chaîne importe le fichier de géolocalisation Insee (format Parquet), associé aux identifiants SIRET, ainsi que les jeux de localisations transmis par les partenaires régionaux.

Source : https://www.data.gouv.fr/datasets/geolocalisation-des-etablissements-du-repertoire-sirene-pour-les-etudes-statistiques

Afin de garantir l'interopérabilité des contributions, les partenaires devront respecter un modèle minimal commun défini par IDéO BFC et présenté en annexe 2.

Ce modèle comprend huit champs obligatoires : SIRET, X, Y, précision, fiabilité, projection, date et partenaire.

Les données partenaires seront transmises au format CSV, encodées en UTF-8, avec une ligne par point de localisation. Les coordonnées pourront être exprimées en Lambert-93 (EPSG : 2154) ou en WGS84 (EPSG : 4326).

Le prestataire devra assurer l'intégration et le contrôle de conformité de ces fichiers, notamment la validité des identifiants SIRET, des coordonnées, des projections et des codes de qualification.

La solution devra également permettre l'intégration de nouvelles sources sans refonte complète de la chaîne de traitement.

2.4 Qualification et sélection

Le prestataire mettra en œuvre la grille régionale de qualification définie par IDéO BFC (annexe 1), en s'appuyant sur les tables de correspondance établies pour les principales sources.

Les traitements devront permettre de :

- Traduire les indicateurs de qualité des sources dans les deux dimensions régionales : précision (P0 à P5) et fiabilité (F0 à F3).

- Conserver les indicateurs de qualité d'origine pour garantir la traçabilité des transformations.
- Comparer les localisations candidates d'un même SIRET selon leur précision, leur fiabilité, leur provenance et leur date de référence.
- Sélectionner une localisation de référence selon des règles explicites et paramétrables, définies avec IDéO BFC.
- Préserver les corrections territoriales fiables et éviter leur remplacement automatique par des localisations moins pertinentes.
- Identifier les situations incohérentes ou non résolues nécessitant un examen complémentaire.
- Déclencher un géocodage complémentaire pour les localisations non satisfaisantes, selon les modalités validées par l'ARNia.

Les règles de correspondance et de sélection devront être paramétrables afin de permettre leur évolution sans modification du code applicatif.

Pour l'étape de géocodage complémentaire, le prestataire privilégiera l'utilisation du service BAN+ de la Géoplateforme, afin d'améliorer la précision des localisations insuffisamment qualifiées.

Il sera également attendu, dans la mesure du possible, un rapprochement avec le Référentiel National des Bâtiments (RNB), afin d'associer un identifiant bâtiment aux établissements localisés et de faciliter les futurs croisements avec d'autres référentiels.

Ces enrichissements devront être intégrés à la chaîne de traitement automatisée, en conservant la traçabilité des sources et des méthodes utilisées.

2.5 Stockage, historique et sélection

La base distingue les localisations sources, la localisation retenue et le journal des changements. Pour chaque sélection, elle conserve au minimum le SIRET, la source retenue, le niveau de qualité, la date du traitement et la règle appliquée. Les anciennes localisations retenues restent consultables.

Le processus est relançable mensuellement, sans doublons ni perte des corrections précédentes. Les échecs d'import ou de traitement doivent être signalés dans un journal exploitable.

La solution devra également gérer les situations dans lesquelles plusieurs points de localisation sont associés à un même SIRET, notamment dans les contributions partenaires.

Elle devra permettre de détecter les doublons de coordonnées, de conserver les localisations distinctes et d'identifier les situations nécessitant un arbitrage.

Aucune fusion, suppression ou sélection arbitraire d'un point ne devra être effectuée sans application d'une règle métier explicite.

Pour chaque localisation candidate, la base conservera les qualifications d'origine et régionales, la provenance, la date de référence et les éventuelles informations de validation.

2.6 Tests et réception

Une première exécution est réalisée sur un échantillon représentatif, puis sur le stock régional. L'ARNia valide les résultats sur un jeu de contrôle comprenant notamment des localisations concordantes, divergentes, absentes et corrigées localement.

La réception intervient après vérification de l'installation, de la bonne exécution de la chaîne, de la conservation des historiques, de la relance mensuelle et de la remise des livrables.

Les tests devront également vérifier la bonne application des correspondances entre les indicateurs de qualité des sources et la grille régionale, la conformité des fichiers partenaires au modèle minimal, ainsi que la gestion des doublons et des localisations multiples.

Un jeu de données test, comprenant notamment des contributions partenaires, sera mis à disposition par l'ARNia pour vérifier ces fonctionnalités.

2.7 Transfert de compétences

Le prestataire organisera une séance de transfert permettant à l'équipe ARNia de comprendre, exploiter et faire évoluer la chaîne de traitement. La séance s'appuie sur la documentation remise.

3. CALENDRIER SOUHAITE

Démarrage dès notification. Durée indicative : 6 à 8 semaines jusqu'à la mise en service, à préciser dans l'offre. Un point de lancement, un point de validation des règles et une recette finale sont prévus.

4. LIVRABLES

Le titulaire fournira au minimum :

- Base PostgreSQL/PostGIS et chaîne de traitement installées et opérationnelles.

- Documentation technique complète et directement exploitable, en français : architecture et schéma de données, sources intégrées, règles de qualification et de sélection fournies par IDéO BFC et leur mise en œuvre, organisation des traitements.
- Documentation d'installation et d'exploitation, en français : prérequis, installation et configuration, lancement et programmation des traitements, ajout ou mise à jour d'une source, modification des paramètres, consultation des journaux, diagnostic des erreurs, sauvegarde et restauration. Elle doit permettre à l'ARNia ou à un tiers de reprendre la solution sans dépendre du prestataire.
- Code source, scripts et fichiers de configuration commentés, ainsi que résultats des tests et compte rendu de recette.
- Modalités d'utilisation des services de géocodage, les règles de rapprochement avec le RNB et les conditions d'attribution des identifiants bâtiments. Le rapprochement avec le RNB ne doit pas être systématique ni forcé. Un établissement localisé uniquement à la rue ou à la commune ne permet pas d'identifier un bâtiment avec suffisamment de certitude.

5. MODALITES DE REPONSE ET DE SELECTION

5.1 Echanges électroniques

Seuls les échanges au travers de la Salle de marchés sont autorisés.

La date limite pour déposer un devis avec une description précise des éléments attendus dans le cahier des charges est **le 01 décembre 2026 à 12h**, aucun dépôt ultérieur ne pourra être pris en compte.

Des questions peuvent être posées par écrit durant la consultation, jusqu'à 5 jours avant la date limite de remise du devis.

5.2 Proposition financière

Le candidat présente un prix forfaitaire HT et TTC, décomposé comme suit :

N°	Désignation	Montant HT
1	Installation et configuration de l'environnement technique	€
2	Intégration des données et développement de la chaîne de traitement	€
3	Tests, recette et mise en production	€
4	Documentation et transfert de compétences	€

Le candidat devra compléter et joindre à son offre l'acte d'engagement ainsi que la décomposition du prix global et forfaitaire fournis dans le dossier de consultation.

Toute prestation optionnelle est chiffrée séparément.

5.3 Proposition technique

Une réponse courte est attendue : compréhension du besoin, méthode et outils proposés avec justification, planning, intervenants et prérequis. Le candidat précise comment il produira la documentation et assurera la reprise autonome de la solution.

Le candidat précisera également, en quelques lignes, les mesures environnementales qu'il prévoit de mettre en œuvre dans le cadre de la prestation.

5.4 Critères de sélection

65 %	Valeur technique
30 %	Prix
5 %	Critère environnemental

La valeur technique apprécie principalement la simplicité, la robustesse, la maintenabilité de la chaîne proposée, ainsi que le recours privilégié à des technologies libres et open source.

Le critère environnemental appréciera les mesures concrètes proposées par le candidat pour limiter l'impact environnemental de la prestation, notamment en matière de sobriété numérique, de dématérialisation des échanges et de limitation des déplacements.

6. CLAUSES ADMINISTRATIVES ET FINANCIERES

En silence du présent cahier des charges, le CCAG-TIC s'applique.

Le règlement de la prestation sera effectué en deux versements :

- 50 % du montant total du marché à sa notification, sous forme d'avance.
- 50 % du montant total du marché à l'issue de la prestation, après réception et validation de l'ensemble des livrables par l'ARNia.