

Société de Calcul Mathématique SA

*Outils d'aide à la décision
depuis 1995*



Accompagnement méthodologique
pour la qualification de données
relatives à la sécurité civile

Rapport Final

adressé au

***Secrétariat Général pour l'Administration
du Ministère de l'Intérieur, région Est***

par la

Société de Calcul Mathématique SA

en application de votre Bon de Commande 1507500179 du 29/09/2016

SCM SA, 17 janvier 2017

Résumé Opérationnel

La plateforme RGE (Référentiel Grande Echelle, où le SGAMI stocke ses données cartographiques) diffuse des données au sein du Ministère, qui souhaite les enrichir en y intégrant des données en Open Data, si elles sont suffisamment pertinentes et de bonne qualité. Les données recherchées sont de type cartographique et portent sur la gestion des crises, les risques naturels, les risques technologiques et les infrastructures réseaux. Le travail demandé était de rechercher des données disponibles en Open Data et de qualifier ces données.

Nos conclusions sur la disponibilité et la qualité des données cartographiques sont les suivantes :

Analyse de la disponibilité des données

La gestion des crises

Des données nationales sont disponibles, mais il s'agit plutôt de bases de données .csv ou Access. Pour les données géolocalisées, le téléchargement de ces données est limité par le nombre d'enregistrements : pour accéder à l'ensemble de ces données, il convient de procéder à plusieurs requêtes, plusieurs sélections. Prenons l'exemple des centres commerciaux : il n'est pas possible de télécharger plus de 50 000 enregistrements à la fois. Il faut procéder à des sélections pour extraire l'ensemble des centres commerciaux sur le territoire français (data.infogreff).

Hormis les thématiques présentées dans le rapport (hôpitaux, centres commerciaux, établissements de tourisme), d'autres thématiques ont été également cherchées au niveau national (écoles, cinémas, lieux de culte, etc.), mais sans succès. Le problème est le suivant : certaines données sont renseignées au niveau régional (mais pas toutes), ce qui rend difficile leur assemblage pour obtenir de l'information au niveau national.

Les risques naturels

Les données liées aux risques naturels sont très rarement disponibles au niveau national. Si c'est le cas, elles sont téléchargeables par département (données BRGM), il convient de procéder à plusieurs téléchargements successifs, ce qui demande un temps non négligeable.

En revanche, nous avons trouvé plusieurs bases de données de très bonne qualité au niveau mondial ou européen : il s'agit de données de la NASA (EARTHDATA) ou l'ESDAC, par exemple. Nous avons pu les télécharger ou les obtenir en écrivant à ces organismes sans trop de problème.

Malgré les recherches approfondies, nous n'avons pas trouvé de données sur les risques d'avalanche et les glissements de terrain.

Les risques technologiques

Des cartes sous le format JPG ou PDF inhérentes aux risques technologiques sont fournies par le site data.gouv. Par ailleurs, nous avons téléchargé la base de données GARSPAR (format Access) qui renseigne, pour chaque commune française, sur le type de risques.

Nous n'avons pas trouvé de données cartographiques inhérentes aux centrales nucléaires, aux sites industriels et aux sites classés Seveso.

En revanche, la base de données Excel IREP, réalisée par la Direction Générale de la Prévention des Risques du Ministère de l'Ecologie, renseigne sur les émissions de polluants par site industriel.

Les infrastructures réseaux

A l'exception du réseau ferroviaire (données SNCF), nous n'avons pas trouvé dans la plate-forme Open Data Soft des données cartographiques inhérentes aux réseaux et infrastructures (réseau de distribution d'eau et de gaz, transport de matières dangereuses, etc.). Nous avons trouvé des cartes au format JPG et PDF par le moyen des sites web des industriels liés à ces activités (RTE, Lyonnaise des Eaux, GDF, etc.) et souvent uniquement au niveau régional ou départemental (Rhône Alpes, Oise, etc.). Ce type d'information est très peu diffusé et peu centralisé au niveau national.

Analyse de la qualité des données

Les données téléchargées sont globalement de bonne qualité : peu de données manquantes ou aberrantes. En revanche, la dernière mise à jour des données est souvent ancienne (2011) et les actualisations sont peu fréquentes. Les métadonnées sont souvent peu, pas ou mal renseignées.

Conclusion générale sur l'Open Data

Peu de données en Open Data sont disponibles pour la France, par rapport aux États-Unis ou à l'Europe. Elles sont souvent fractionnées, c'est-à-dire récupérables par département et non à l'échelle nationale. Les données sont globalement de bonne qualité mais les métadonnées sont souvent peu, pas ou mal renseignées. L'Open Data français a encore des progrès à faire.

Les données cartographiques que nous avons rencontrées existent sous différents formats (Shape, Raster, etc.) et avec des résolutions différentes, ce qui rend difficile un croisement des données. Exemple : des données sur les zones d'inondation (format Raster, résolution 100 m) et sur les zones sismiques (format Raster, résolution 1 km) seront difficiles à croiser.

Le Ministère de l'Intérieur pourrait recommander à l'ensemble de ses partenaires travaillant dans l'Open Data un format commun et une résolution commune pour améliorer la cohérence de l'ensemble des données cartographiques.