

# QUELQUES PRÉCAUTIONS D'USAGE DES SIG

*Serge Motet  
Ecole Nationale des Sciences Géographiques*

Au moins chez les enfants, il y a dans toute chose nouvelle, une période d'engouement suivie d'une période de désintérêt. Et il est plus utile de prévenir la désaffection que de susciter l'intérêt.

Cela s'applique au système d'information géographique. Il est clair que l'intérêt pour les SIG existe et se développe grâce au dynamisme des chercheurs, des prestataires de services et des distributeurs de logiciels.

Cependant, au-delà de l'attrait de la nouveauté, les expériences précédentes incitent à la prudence.

Les références dans d'autres domaines sont nombreuses et instructives. Un SIG est avant tout un système d'information, et on peut prévoir que son évolution suivra celle des autres systèmes d'information. Même si ces derniers concernent des domaines aussi éloignés de la topographie que la gestion ou la CAO automobile.

A ce titre, deux constats sont à la fois déterminants et antagonistes :

- D'une part, le développement des systèmes d'information numérique est inéluctable. Ces systèmes s'imposent dans tous les domaines de l'activité économique.
- Mais d'autre part, ce développement est difficile à cause de l'inadaptation des solutions informatiques et du manque de prévision sur les besoins et sur les effets des systèmes. En conséquence, certaines entreprises ont perdu beaucoup de temps et d'argent dans un système d'information mal adapté.

Utiliser l'information numérique, c'est à la fois indispensable et risqué. Cette opposition s'applique aux SIG et elle pose problème aux utilisateurs d'information géographique. A tel point qu'un guide du gouvernement du Québec (CDPGG) sur la géomatique n'hésite pas à parler du stress du gestionnaire du projet de SIG.

Les expériences précédentes montrent que la difficulté réside dans l'évaluation du risque. Le risque apparaît à trois étapes :

Au départ, dans le choix de l'organisation et des spécifications, le risque est à la fois économique (adéquation des objectifs aux besoins) et technique (choix des solutions adaptées aux objectifs).

Puis, lors de la saisie, la difficulté réside dans l'incertitude d'obtenir la qualité requise.

Enfin, au cours de l'entretien du système, il faut s'assurer de la qualité de la mise à jour.

On ne peut malheureusement donner de solutions générales contre l'incertitude. Cela dépend des applications. Mais il me semble nécessaire de préciser les sources de problèmes.

## DÉFINITION D'UN SIG

Au préalable, il est utile de définir le terme SIG car il existe au moins deux interprétations. Pour la première, un SIG est un ensemble de logiciels pour la saisie, le stockage, l'analyse et la représentation de l'information géographique numérique. La deuxième définition est plus vaste : un SIG comprend non seulement les logiciels mais aussi des données sur une zone ou un thème particulier. On parlera du SIG régional du Languedoc-Roussillon pour nommer un ensemble d'outils et de données mis en place dans cette région.

C'est cette dernière définition qui est utilisée ici. Elle est illustrée par la figure 1 issue de (ARONOFF).

## JUSTIFICATION ÉCONOMIQUE

La rentabilité d'un projet est la première condition de son existence. Il y a deux visions de la rentabilité :

D'une part, il y a l'objectif de rentabilité des prestataires de services en SIG. Cette activité s'est développée d'abord sur le continent Nord-Américain et est en train de s'étendre en Europe. On peut distinguer les domaines suivants :

- La saisie et les activités connexes : définition du processus de saisie en fonction des besoins, contrôle de la qualité,...
- Le conseil en SIG qui va de la comparaison de logiciels jusqu'au montage de base de données localisées.
- Le développement de logiciels.
- La cartographie et la PAO.

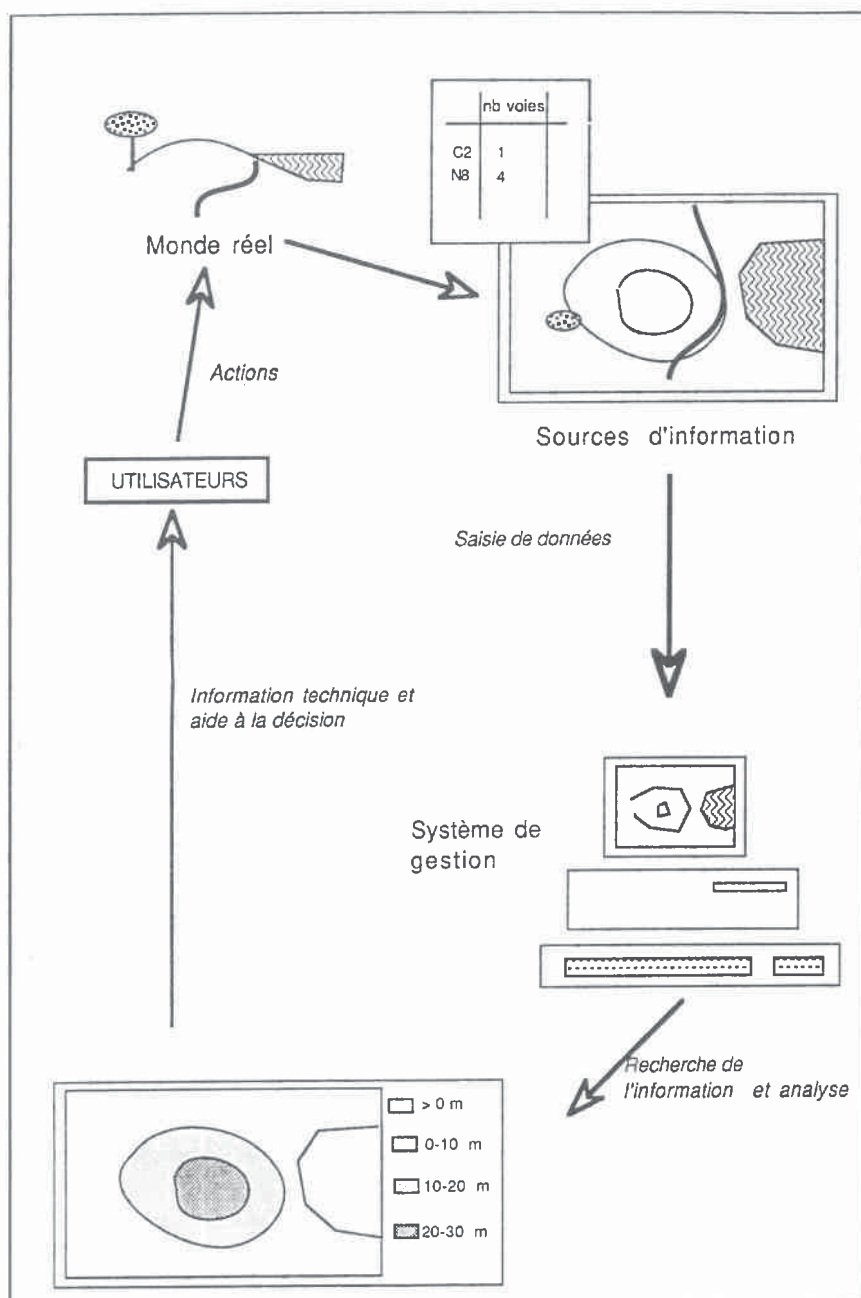


Figure 1.

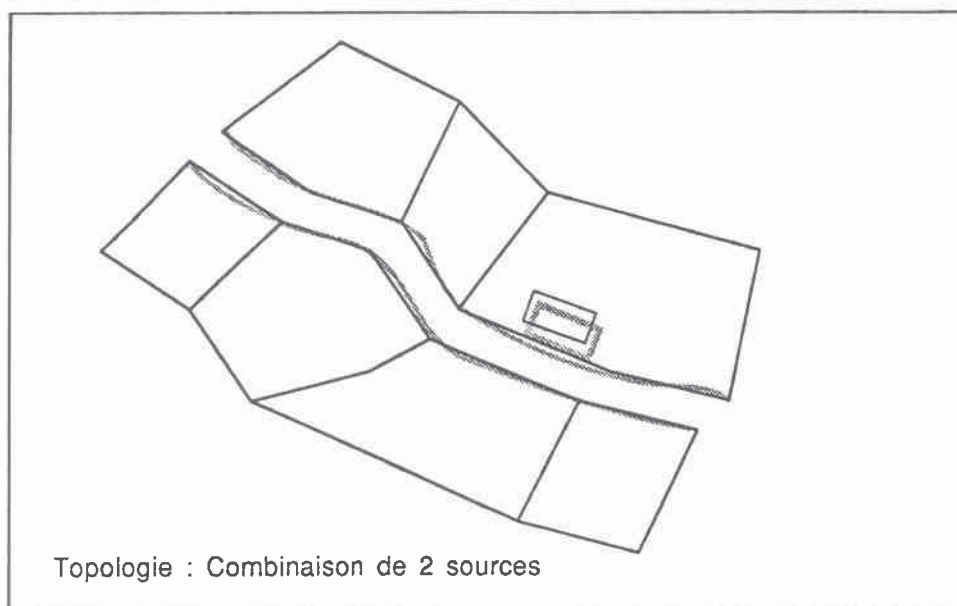


Figure 2.

D'autre part, il y a le point de vue des gestionnaires d'information géographique. Il s'agit de certains services des collectivités territoriales, des gestionnaires de réseaux ou des ministères.

Le problème dans ce cas est d'évaluer les bénéfices engendrés par un SIG. Il faut tenir compte de tous les gains directs ou indirects. Un certain nombre de méthodes sont présentées dans (DIDIER).

## TOPOLOGIE

Les difficultés posées par la topologie sont peut-être les plus difficiles à cerner.

On peut définir une information topologique comme une information sur la superposition ou la mitoyenneté de deux détails topographiques. Par exemple, un carrefour entre deux routes, la façade d'une maison sur une place sont des notions topologiques qui rendent compte de l'intersection des routes ou de la mitoyenneté de la place et du bâtiment.

Ces notions sont indispensables pour les applications. Par exemple, la connaissance de la géométrie des routes dans sa topologie - c'est-à-dire le réseau - présente un intérêt limité. Elle ne permet pas les calculs sur les itinéraires ou sur la répartition des centres de distribution.

Face à ces avantages, la topologie génère de nombreuses complications. Par exemple, il est plus difficile, lors de la saisie, de maintenir la topologie lorsque l'information provient de plusieurs plans. Même si chaque plan a une topo-

logie correcte, la topologie entre détails de plans différents est souvent faussée à cause de l'imprécision géométrique.

Cela est illustré par la figure 2 où deux plans sont représentés : l'un en trait plein et l'autre en grisé. L'erreur se situe sur le bâtiment rectangulaire. On voit que sa représentation en grisé empiète sur la route en trait plein (ce qui est contraire à la contrainte topologique selon laquelle les emprises d'une route et d'un bâtiment ne se recouvrent pas).

Ce type de difficulté ne peut se résoudre que par des consignes de saisie précises. Ce qui suppose au préalable une étude complète des sources et de la qualité souhaitée.

## QUALITÉ

La qualité de l'information gérée dans un SIG détermine sa capacité à répondre aux besoins. C'est donc un facteur prépondérant.

La qualité de l'information géographique pose un problème de définition et de mesure. C'est encore un sujet de recherche. Mais, autour de certaines notions, apparaît maintenant un consensus.

Ces notions ont reçu des définitions précises mais elles présentent le défaut d'être interdépendantes.

On peut distinguer trois niveaux de qualité :

- Le niveau local concerne des mesures de qualité propres aux détails. Il s'agit de la précision géométrique ou sémantique, de la résolution et de la cohérence logique. La précision géométrique est l'écart entre la position des entités numériques par rapport aux détails réels. C'est une notion connue en topographie, qui s'exprime par le biais, l'erreur quadratique moyenne, etc. Elle se mesure à partir d'échantillons. La résolution est l'aire ou la longueur minimale des entités représentées. Elle est liée à la nature du détail. Elle peut être contrôlée de manière automatique.

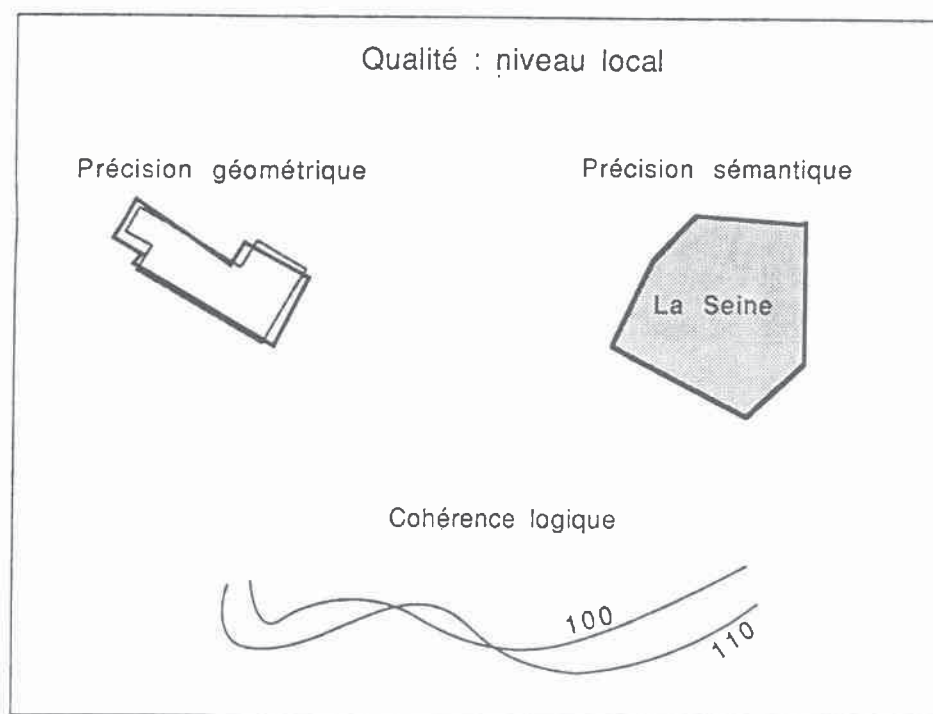


Figure 3.

La précision sémantique ou précision d'attribut exprime l'exactitude de la valeur des attributs non géométriques. C'est par exemple la mauvaise orthographe d'un toponyme. Les mesures sont soit l'écart à la valeur nominale pour un attribut à valeur continue, soit le pourcentage de valeurs fausses pour les attributs dont le nombre de valeurs possibles est fini. Les méthodes de mesure sont fondées sur l'échantillonnage ou sur la recherche automatique des anomalies. La cohérence logique est le respect des contraintes et des spécifications. Les exemples d'erreurs sont des intersections entre courbes de niveau. Cette notion contient la résolution. On peut la mesurer par des contrôles automatiques.

- Au niveau du lot de données, on a une vue globale de la qualité des données. Il comprend l'exhaustivité, l'actualité et la généalogie. L'exhaustivité mesure l'excès ou le manque d'information. Elle rend compte des entités manquantes ou fictives. Elle s'exprime par un pourcentage d'oubli ou d'ajout. La mesure est difficile. Elle peut être faite par une vérification systématique sur une zone-test. L'actualité est le pourcentage de changement entre la date de la collecte et le présent. Elle peut s'évaluer à partir de la date des documents et d'un taux annuel d'évolution. La généalogie est l'ensemble des informations qui permettent de retrouver l'origine et les traitements effectués sur le lot de données. La généalogie est indispensable pour rattraper les erreurs, prévoir la mise à jour et évaluer la qualité du niveau local.
- Le niveau utilisateur se situe à la limite de la qualité. Certains auteurs regroupent à ce niveau le coût et l'accessibilité. Dans l'acces-

sibilité, on distingue principalement la normalisation du modèle de données et la disponibilité dans des formats d'échange usités.

Avant de clore ce chapitre sur la qualité, il faut insister sur la difficulté d'assurer une qualité a priori. En effet, il est simple, mais parfois coûteux, de contrôler la qualité d'un lot de données déjà saisi. Mais il est plus délicat de définir un processus de saisie en vue d'obtenir une qualité donnée.

## CONCLUSION

On me reprochera probablement d'avoir montré les embûches sans apporter de solutions. J'ai donné la liste des domaines économiques, mais quels sont les plus intéressants ? J'ai évoqué les problèmes liés à la topologie, mais comment les résoudre ? J'ai défini la qualité, mais comment l'atteindre ?

Il n'y a pas de réponse simple et générale à ces questions. Pour réussir un SIG, il faut trois ingrédients :

- **De la formation.** On doit apprendre non seulement l'utilisation des logiciels, mais aussi d'autres domaines comme l'évaluation de la qualité d'une source d'information.
- **De la méthode.** Les SIG n'ont pas encore leur méthode Merise, et cela manque. Cependant on peut s'inspirer de méthodes existantes.
- **De l'expérimentation.** A mon avis, on ne prend conscience des difficultés qu'après avoir essayé de constituer un SIG. Il faut accumuler une certaine expérience. Et la faire partager. On peut donc souhaiter que ces colonnes permettent de rassembler d'autres expériences en SIG.

## BIBLIOGRAPHIE

(CDPGG) - Comité directeur du plan géomatique gouvernemental : Guide de gestion ; introduction de la géomatique au gouvernement du Québec. Québec, Canada. Mai 1990.

(ARONOFF) - Stan Aronoff : Geographic Information Systems ; A Management Perspective. WDL Publications. Ottawa, Canada. 1989.

(DIDIER) - Michel Didier : Utilité et valeur de l'information géographique. Economica. Paris. 1990.

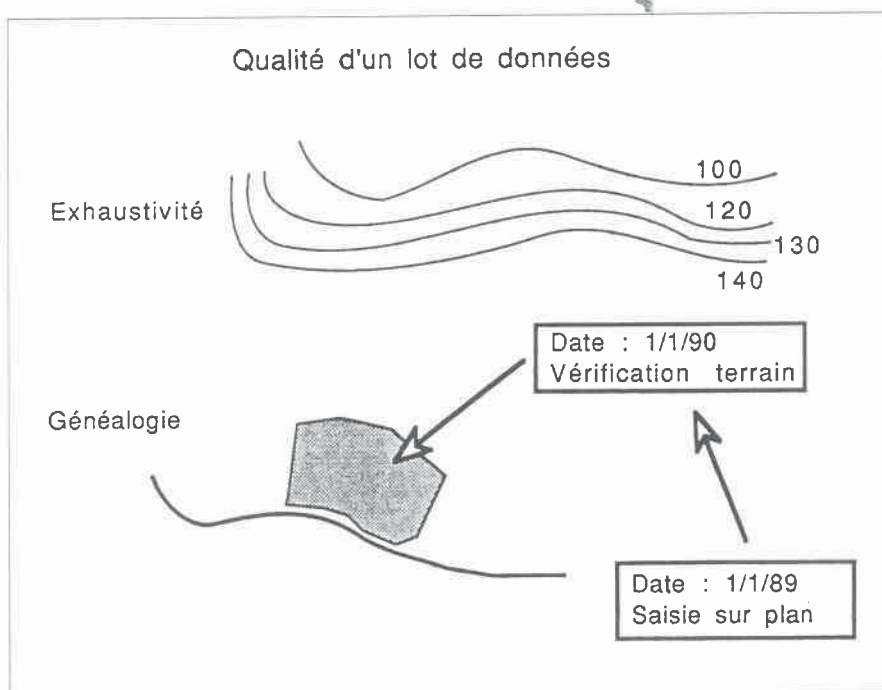


Figure 4.