

# Qualité des données : évaluer ou juger ?

AXES Conseil



Pôle Systèmes d'information de Ségit

Ingénierie  
axes.fr

axesconseil@axes.fr



Dans le cadre de ses activités, AXES Conseil réalise fréquemment des analyses qualitatives de données SIG, notamment orientées vers les réseaux humides. A ce jour, plus d'une trentaine d'analyses a été réalisée.

La présentation de l'analyse provoque des réactions variables en fonction du mode de production de la donnée et du niveau de qualité observé.

## Un accueil de l'évaluation variable selon le contexte et le résultat

Dans le cas d'une production externe, l'analyse des données est accueillie avec intérêt, car elle permet d'évaluer objectivement le travail d'un prestataire ou délégataire. Le travail de correction éventuel sera réalisé par ce dernier.

Lorsqu'il s'agit d'une production interne, les retours sont plus contrastés : Si, dans la grande majorité des cas, un intérêt marqué prédomine, il arrive dans certains cas, des utilisateurs montrent un intérêt moindre. La nature de la réaction repose notamment sur les résultats obtenus :

- quand les résultats sont globalement bons, l'entité y porte un intérêt significatif car l'analyse vient « valider » la qualité du travail réalisé en interne : Cela est gratifiant pour les producteurs.
- quand les résultats révèlent des faiblesses vis-à-vis des spécifications ou des attendus d'usage, certains producteurs peuvent chercher minimiser les constats sur les résultats : « *Les utilisateurs ne se plaignent pas* » ; « *Cela est suffisant pour nos usages* » ; « *Nous n'avons pas le temps / les ressources pour mieux faire* » ; etc. Pouvant se sentir jugés sur leur travail, des producteurs peuvent chercher à justifier la situation par les motifs externes, pouvant être parfois bien réels.

## La qualité est liée aux usages

Il est important de rappeler que la qualité est « **la capacité du produit (exemple : donnée), service ou système à satisfaire les exigences spécifiées** » (source : iso.org).

La qualité est donc définie comme la concordance entre des usages attendus (spécifications) et le niveau de satisfaction de l'utilisateur (réponse aux spécifications).

Une donnée est donc de bonne qualité quand elle répond aux besoins d'usage de l'utilisateur. **Il n'existe pas de qualité de données absolue.** Les résultats d'une analyse qualité doivent donc être systématiquement contextualisés avec les usages utilisateurs et leur satisfaction.

### Illustration

*Une analyse qualitative d'une donnée modélisant un réseau d'assainissement a montré :*

*>Un très bon niveau de complétude : nombre d'objets et renseignements attributaires,*

*>Un faible niveau dans les relations topologiques des canalisations*

*Les agents de terrain impriment des plans papiers : ils considèrent la donnée de bonne qualité car l'ensemble des objets est présent.*

*Les agents travaillant sur la modélisation rencontrent des problèmes sur les parcours de graphe : ils considèrent la donnée de mauvaise qualité.*

## Analyser une donnée, ce n'est pas juger son producteur...

L'analyse de la qualité d'une donnée ne doit pas être perçue comme un jugement sur le travail du producteur, mais comme une évaluation de l'adéquation de la donnée avec les usages prévus. Or, si les usages évoluent dans le temps, les modèles de données ou les modes de productions ne sont pas toujours revus ou adaptés en conséquence. Dans le temps, des décalages peuvent donc se faire jour entre usages et exigence sur les données.

**... mais fixer un point de départ d'actions d'amélioration ou d'exploitation.**

L'objectif d'une analyse qualité d'une donnée serait donc d'identifier ses faiblesses vis-à-vis des usages courants ou identifiés à l'avenir (contexte IA notamment).

Au-delà des éléments techniques, une analyse de la qualité des données fournit des arguments objectifs qui permettent de sensibiliser les décideurs et de planifier des actions, qu'elles soient ponctuelles (corrections, mises à niveau) ou dans le long terme (mise en place d'outils de contrôle et de suivi).