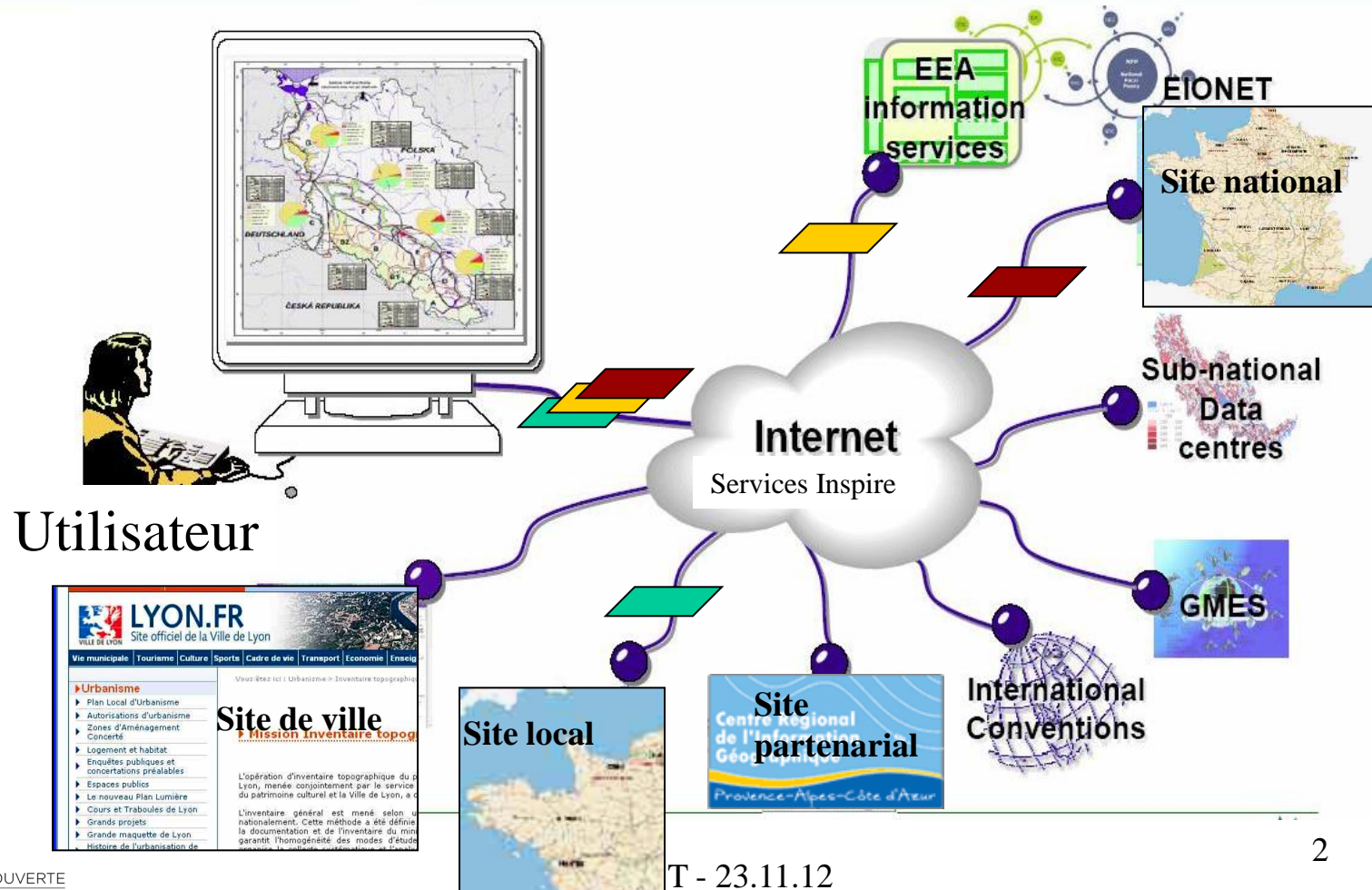


# Données géographiques : accessibilité et cohérence

Marc Leobet

Mission de l'information géographique  
CGDD/DRI


# L'enjeu, assembler des sources diverses pour une utilisation donnée



# Un constat sur la cohérence géométrique



# Les constats sur l'utilisabilité

- 
- la rupture des relations entre producteur et utilisateur
  - Une difficulté commune au Big data et à l'Open data
  - une influence directe sur :
    - la fiabilité des analyses spatiales produites
    - et les décisions qui en découlent

# Quelle cohérence avec quoi?

- Le règne de la donnée empirique :
  - De très nombreux producteurs;
  - Peu ou pas de spécifications de données;
  - Une production personnalisée par producteur
- Un fonctionnement *hors du monde contrôlé* :
  - Principe de réutilisation par des inconnus pour des usages inimaginables
  - Principe de coopération (ajouts, retraits)

# Les attentes des autorités

- Les attentes de la Commission européenne sont difficiles à cerner.
- Les attentes françaises peuvent être résumées par :
- "n'importe quelle donnée vaut mieux que pas de données".
- L'enjeu majeur : la réutilisation des données publiques,
- La donnée n'est plus conçue pour correspondre au besoin du réutilisateur, mais celui-ci doit pouvoir évaluer cette correspondance
- Un enjeu particulier dans la précision sémantique

# Des solutions écartées

- Inadéquation des solutions à base ISO (19113 + 19114)
- Opposition totale avec le concept de certification des données par un tiers.
- La conformité n'est pas synonyme de qualité externe
- La notion de conformité en interrogation
  - Et si on est conforme à 90% (avec des spécifications) ?
- Un manque certain de maturité de ces aspects

# Des solutions envisagées

- Un outillage le plus simple possible;
- Un objectif de coordination, facilité par la proximité territoriale des producteurs;
- Un enjeu de cohérence sémantique;
- Un besoin de description de la donnée
  - Donner les clés à l'utilisateur
- Importance-pivot du champ généalogie
  - Pour comprendre par le processus les limites de la donnée.



# L'outil des métadonnées

- **Intitulé** du jeu de données : son titre
- **Résumé** : à quoi répond-il?
- **Mots-clés** : pour les moteurs de recherche
- **Date** : pour la validité de la donnée
- **Généalogie** : principale information pour estimer la qualité
- **Résolution** : échelle d'emploi privilégiée
- **Restrictions concernant l'accès public**
- **Spécifications de référence**
- **Conformité** : oui / non / non évalué

# Conclusion

- L'accessibilité des données publiques est lancée et croîtra;
- Leur cohérence ou leur qualité n'est pas un pré-requis;
- La qualité pour quels utilisateurs ? Pour combien de type d'utilisateurs ?
- Il faut poursuivre les études et gagner en expérience pour mieux définir les bonnes pratiques
- Les décisions réglementaires ne viendront qu'à long-terme.