

DATAVIZ ET PLATEFORMES DE DONNÉES TERRITORIALES

Sommaire

Dataviz et plateformes de données territoriales	1
Article 1 - Pourquoi ?	2
Article 2 - Quels besoins ? Quels usages ?	4
Article 3 - L'expérience de la directive INSPIRE	7
Article 4 - Quelle voie suivre ?	9

Préambule

Dans le cadre de l'animation du Réseau des CRIGEs, l'Afigéo mène une veille active sur les stratégies data dans les territoires et sur l'actualité et les enjeux prospectifs sur les usages de la géomatique. Elle a produit en 2023 une publication "[Plateformes territoriales de données; enjeux et perspectives](#)" avec le soutien du Ministère de la Transition Écologique et de la Cohésion des Territoires (MTECT) et propose, au travers de ce recueil d'articles, d'explorer les opportunités et défi proposés par la valorisation des données géographiques au sein des plateformes de diffusion de ces mêmes données.

Production : Afigéo / Benjamin CHARTIER (Optéos), avec le soutien de l'Ecolab du MTECT, -
Juin 2023

Article 1 - Dataviz et plateformes de données territoriales – Pourquoi ?

La dataviz a connu un essor fulgurant ces dernières années. On la rencontre maintenant quasiment quotidiennement dans nos milieux professionnels mais aussi dans la presse grand public et dans nos applis mobiles. Elle est devenue un support de communication inévitable dès lors que l'on cherche à transmettre ou percevoir des phénomènes de manière synthétique à partir de données souvent beaucoup plus riches en détails et complexes. La dataviz a démontré sa capacité à répondre aux attentes à la fois d'un public pressé et d'un public ne disposant pas nécessairement des compétences nécessaires à l'analyse de telles données.

Il n'est donc pas étonnant que la plupart des plateformes de données territoriales positionnées initialement sur la question de la diffusion des données géographiques (avec la mise en œuvre de la directive européenne INSPIRE notamment) se questionnent sur la place à donner à la dataviz. J'attire votre attention sur le fait qu'il ne s'agit pas, selon moi, juste d'un positionnement opportuniste par rapport à un phénomène que certains seraient tentés de qualifier de mode passagère. Je vous explique...

Tout d'abord, il existe un parallèle évident entre dataviz et cartographie. La cartographie, dans la plupart des cas (dès lors qu'elle repose sur des données) n'est rien d'autre qu'une dataviz. Une dataviz qui repose sur des principes un peu particuliers mais une dataviz tout de même¹. Partant de ce principe, et constatant que de nombreuses plateformes de données territoriales préparent ou ont réalisé leur mutation de la donnée purement géographique à la donnée en général, il paraît évident qu'une plateforme qui donne à voir des cartes devrait aussi laisser une place à la dataviz. Vous trouverez cela d'autant plus logique si vous admettez que la cartographie n'est pas toujours l'approche la plus adaptée même pour des données qui ont une composante géographique.

Ensuite, vous avez peut-être remarqué que j'ai utilisé le terme « parallèle » pour comparer la place de la dataviz et de la cartographie. Votre cerveau a peut-être marqué un moment d'arrêt en le lisant. C'est sans doute que ce terme n'est pas idéal. Le terme « intrication » serait plus adapté. En effet, il est plutôt courant d'observer des graphiques intégrés dans des cartes et de trouver des cartes dans des tableaux de bord. On constate également que dans le bestiaire riche de la dataviz on trouve des représentations graphiques proches de la carte géographique sans pour autant s'appuyer sur des dimensions spatiales. Même si les outils peuvent être sensiblement différents, ces deux domaines partagent quelques points communs notables : des concepts, des méthodes (la sémiologie), et surtout, des données, des utilisateurs et des besoins.

¹ Voir en particulier les travaux de Emile Cheysson à la fin du 19ème siècle:
<https://www.icem7.fr/datavisualisation/emile-cheysson-et-lage-dor-de-la-datavisualisation-francaise/>



Et c'est surtout là que je voulais en arriver. On peut être effrayé par les différentes techniques et graphiques de la dataviz par rapport à la cartographie : pour la plupart des plateformes de données territoriales, maîtriser ce domaine peut être un réel défi compte tenu des compétences que cela peut requérir. Ceci dit, il faut bien constater que la raison d'être de nombreuses plateformes réside dans leurs données, leurs utilisateurs et leurs besoins. Or, même si produire une dataviz et une cartographie ne relève pas des mêmes compétences techniques, ces représentations graphiques peuvent s'appuyer sur le même patrimoine de données, s'adresser aux mêmes utilisateurs et répondre aux mêmes besoins. Ces plateformes de données territoriales généralistes ont souvent des objectifs qui vont au-delà de la simple mise à disposition ou diffusion des données de leurs membres. Elles valorisent l'écosystème de la donnée dans son ensemble (données, outils, producteurs, utilisateurs...) et facilitent leurs réutilisations.

Puisque nous évoquons la valorisation, nous pouvons noter que l'avènement des plateformes de données ouvertes a banalisé quelque chose que nous n'observions pas jusque-là sur les plateformes de données géographiques : la mise en avant des réutilisations des données. Or, la dataviz est un type de réutilisation dont le pouvoir de communication est très efficace. C'est aussi pour cela que la dataviz ne devrait pas être négligée sur les plateformes de données territoriales actuelles.

La dataviz, au même titre que la cartographie géographique, répond parfaitement à ces objectifs de valorisation et de facilitation. Elle devrait donc occuper une place similaire au sein des plateformes de données territoriales et devrait donc augmenter la portée de leurs actions conformément à leurs objectifs existants.

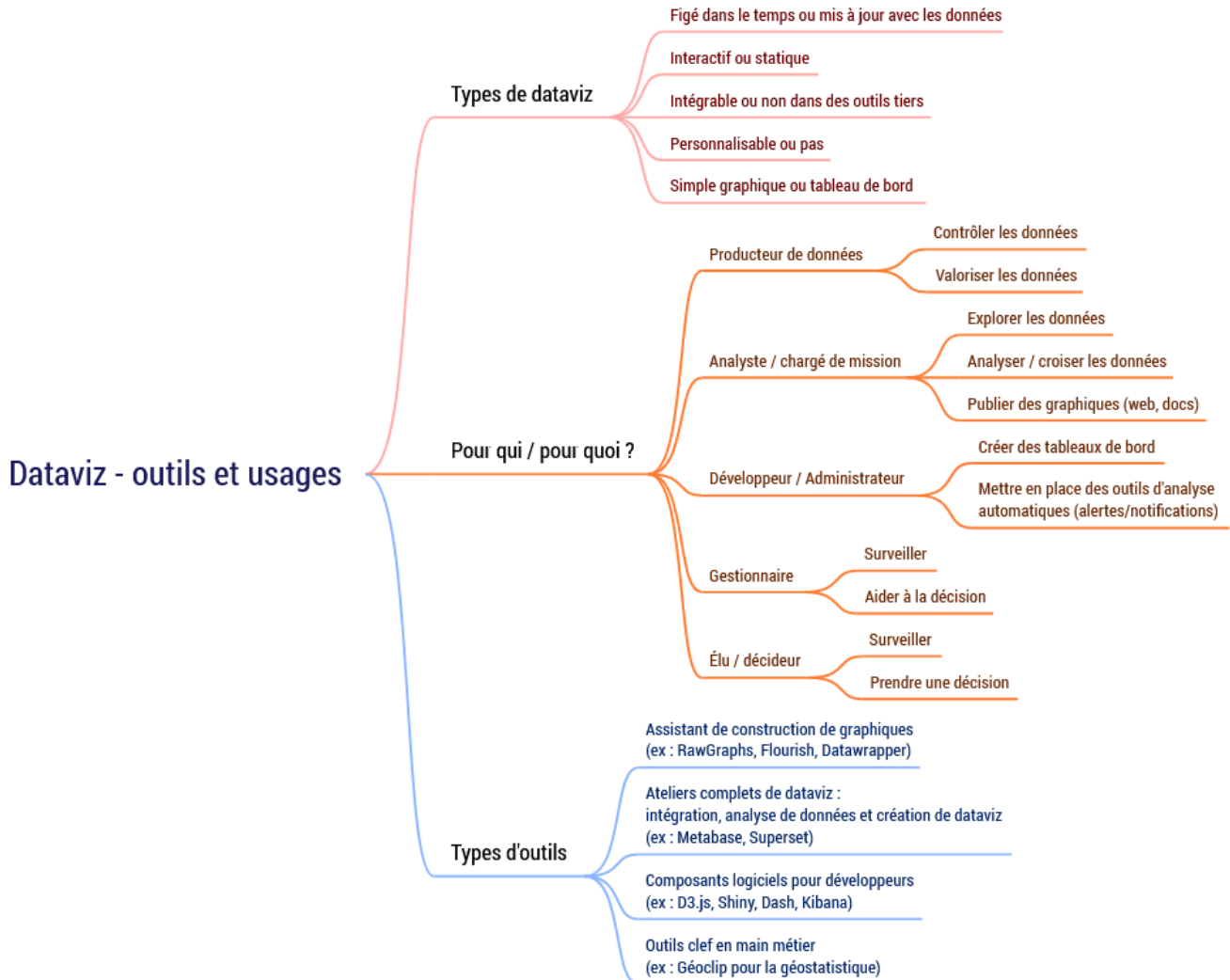
Article 2 - Dataviz et plateformes de données territoriales – Quels besoins ? Quels usages ?

On l'a vu précédemment, la place de la dataviz au sein d'une plateforme de données est tout aussi légitime que celle de la cartographie. Mais, il faut peut-être revenir aux fondamentaux : la dataviz pour qui et pour quoi faire, finalement ?

Le besoin que l'on entend souvent émerger est assez simple : « J'ai des données ; j'ai besoin d'un outil pour en faire des dataviz. ». La simplicité de l'expression d'une demande cache bien souvent une complexité que son auteur ne soupçonne pas. Non pas qu'il soit incompréhensible mais parce qu'elle peut cacher des situations bien différentes et difficiles à concilier. Pour illustrer cela, voici quelques exemples de dataviz que les utilisateurs et les administrateurs des plateformes de données trouveraient intéressants de produire :

- Un graphique exploratoire d'un jeu de données permettant à un usager de mieux comprendre sa structure, d'en évaluer la qualité et de se faire une idée de son adéquation vis-à-vis d'une réutilisation qu'il souhaite réaliser.
- Une infographie au design soigné qu'un réalisateur pourra intégrer à un autre support.
- Un outil interactif permettant à un usager de personnaliser la représentation d'un jeu de données afin de l'intégrer sur un site web par exemple ou simplement de le partager avec ses interlocuteurs.
- Un tableau de bord thématique croisant plusieurs sources de données évoluant dans le temps donnant un état des lieux en temps réel. Exemple : surveillance de crues de cours d'eau exploitant des données météo, des données de capteurs, des modèles de simulation, des données sur les infrastructures, les biens et les personnes éventuellement menacés. Il s'agit ici d'un outil décisionnel conçu par des spécialistes et mis à disposition d'opérateurs ayant des responsabilités (prise de décision).

Certes, tous ces exemples sont des dataviz mais leurs natures, leurs modes de production, les possibilités qu'elles offrent, les utilisateurs cibles sont très différents. Pour aider à mieux percevoir cette variété et donc la complexité à répondre à un besoin qui paraît simple vous trouverez ci-dessous une carte mentale qui identifie quelques pistes de réflexion sur l'usage de la dataviz au sein des plateformes de données territoriales.



Si vous ajoutez à cela l'extraordinaire fertilité du bestiaire de dataviz (voir illustration ci-dessous), vous comprendrez alors aisément qu'il n'existe pas de solution idéale pour les plateformes de données territoriales en matière de dataviz.



Exemples de graphiques issus de [Data Viz Project](#) et [son catalogue](#).



Article 3 - Dataviz et plateformes de données territoriales – L'expérience de la directive INSPIRE

Vous l'aurez compris en lisant l'article précédent : essayer de répondre à toutes les attentes en matière de dataviz est extrêmement difficile tant la variété de ses formes et de ses usages est importante. Et cette difficulté se ressent particulièrement pour les plateformes de données territoriales dans la mesure où, structurellement, elles ne peuvent pas s'appuyer sur un cadre homogène d'emploi des données :

- Elles ne sont en général pas focalisées sur un type de données particulier ni sur une thématique spécifique ;
- Elles se sont en général structurées de manière partenariale en fédérant des usagers de natures très hétérogènes.

Il devient évident que ces plateformes peuvent difficilement répondre à l'ensemble des besoins de leurs usagers, et sûrement pas à l'aide d'un outil unique et idéal. Nous pouvons évidemment nous en douter puisque cette histoire s'est déjà déroulée au cours de la mise en place des plateformes de données géographiques territoriales pour la mise en place d'outils de cartographie. Cette situation est en tout point similaire en termes de variété de données, de besoins et d'usages. Elle a débouché sur une situation intéressante :

- La mise en place par les plateformes de données elles-mêmes d'applications web relativement simples et efficaces dans le sens où elles offrent des fonctions peu nombreuses et à faible valeur ajoutée offrant, dans la grande majorité, un service minimum sur une grande variété de données disponibles en ligne. Ces fonctions sont souvent de l'ordre de la composition simple d'une carte, de la prévisualisation des données, du filtrage attributaire ou du téléchargement des données. Il s'agit de mettre à disposition des outils pour les premiers usages des données, voire d'outils qui permettent aux plateformes de valoriser leurs données et de donner envie d'aller plus loin avec elles.
- La persistance d'outils SIG bureautiques classiques (à l'image de QGIS et d'ArcGIS par exemple) alors que certaines prédictions émises il y a une dizaine ou quinzaine d'années imaginaient leur remplacement à terme par des solutions complètement en ligne.
- L'émergence de solutions indépendantes capables de se connecter aux API d'accès aux données proposées par les plateformes de données pour proposer des services spécifiques.

En somme, nous avons vu s'installer un écosystème d'outils développé plus ou moins consciemment et capable de fournir des réponses à des attentes et besoins variés. Il n'est pas toujours simple de se retrouver dans cet environnement d'outils et de données hétéroclites. Cette situation n'est pas parfaite mais elle a l'avantage de permettre une



certaine agilité et une forme de résilience. Je veux dire par là que n'importe quel outil technique déjà en place qui viendrait à être obsolète verrait sans doute sa place prise par un autre sans probablement que cela ne provoque de catastrophe industrielle. D'un côté cette situation est imparfaite car elle demande aux usagers des plateformes de données territoriales de faire preuve d'agilité en jonglant avec de nombreux outils. D'un autre côté, je ne perçois aucun signal qui montrerait une tendance inverse (à savoir la mise en place d'outils numériques peu nombreux et concentrant la quasi intégralité des fonctionnalités dont ont besoin les géomaticiens et cartographes). Voilà donc une situation propre aux plateformes de données territoriales dont on peut se plaindre au quotidien mais qui ne semble pas prête d'évoluer dans les grandes lignes faute de meilleure alternative.

L'intérêt de cette situation réside dans les raisons qui l'ont amenée. En premier lieu, vous noterez que les plateformes de données géographiques territoriales ont été créées dans un esprit partenarial, mutualiste et souvent contributif, et qu'elles se sont construites de manière progressive en essayant de répondre à de nouveaux besoins au fur et à mesure. Il est clair que, dans un tel cadre, la manière la plus efficace de satisfaire une grande variété de besoins n'est pas de concevoir un système monolithique qui s'impose à tous de manière uniforme. Par contre, la modularité et l'interopérabilité sont de bonnes cartes à jouer. Si l'on ajoute à cela un cadre réglementaire (directive européenne INSPIRE) qui impose certains principes d'interopérabilité et qui promeut des standards clairement identifiés (à la différence de la Loi pour une République Numérique, qui encourage l'ouverture des données), on comprend mieux comment un tel écosystème moyennement satisfaisant de solutions hétérogènes disposant de périmètres fonctionnels plutôt limités peut prendre place sans que l'on perçoive de réelle alternative.

Article 4 - Dataviz et plateformes de données territoriales – Quelle voie suivre ?

Quelle sorte d'enseignement peut-on en tirer pour la dataviz ? Même si les domaines de l'information géographique et de la dataviz ont beaucoup de points en commun, nous pouvons noter d'importantes différences qui ont un impact majeur sur les aspects techniques des solutions.

Premièrement, la dataviz ne bénéficie pas d'une impulsion réglementaire telle que la directive INSPIRE. Celle-ci avait tracé la voie pour ce qui concerne des fonctionnalités minimales à implémenter au niveau des plateformes de données et avait permis de définir un noyau commun de protocoles de communication.

Deuxièmement, les questions d'interopérabilité et de modularité ne semblent pas être prioritaires pour les éditeurs de logiciels de dataviz. On trouve de nombreuses solutions monolithiques, qui favorisent l'intégration de données dans leur propre base au préalable plutôt que leur accès distant. Certaines solutions permettent d'intégrer une URL d'accès à un fichier JSON ou CSV mais la plupart du temps la réplication des données dans la base de données propre de l'outil de dataviz est privilégiée. Le marché semble dominé par des solutions qui ne favorisent pas l'interaction d'outils hétérogènes.

À titre d'exemple vous pouvez consulter la captation d'un webinaire de DataGrandEst portant sur l'usage de plusieurs outils de dataviz qui illustre bien les limites actuelles de ces solutions dans le cadre des plateformes de données partenariales :

[Création de tableaux de bord, outils, techniques et astuces !](#)

En matière d'interopérabilité et de dataviz notons toutefois quelques exceptions même si elles sont extrêmement spécifiques :

- Les projets Vega et Vega-Lite qui définissent un langage de définition de graphiques (Vega is a visualization grammar, a declarative language for creating, saving, and sharing interactive visualization designs). Ce langage peut être notamment utilisé par Flourish et Altair.
- Les protocoles TJS² (qui n'a réellement été adopté que par Géoclip) et SDMX (adopté par les instituts statistiques nationaux et européens).

Troisièmement, de manière contre-intuitive, réaliser une dataviz à partir de données présentes sur les plateformes de données territoriales n'est pas aussi aisé que de réaliser une cartographie. Il serait tentant de penser que la composante géographique est tellement spécialisée qu'elle serait complexe à intégrer dans une cartographie alors que des données

² Pour une introduction sur ce standard, voir ce billet de 2011:

<https://blog.georezo.net/geointerop/2011/10/04/tjs-le-nouveau-standard-de-partage-de-donnees-geo-statistiques/>



quelconques ne posent aucune difficulté dans le cadre de la production de dataviz. Ce serait oublier que la plupart de ces graphiques s'appuient sur des classifications ou agrégations selon 2 dimensions distinctes. La grande majorité des jeux de données disponibles actuellement sur les plateformes de données territoriales n'offrent pas cette possibilité soit parce qu'ils n'ont pas été conçus pour cela soit parce qu'ils nécessitent d'être croisés avec d'autres jeux de données. Cette difficulté à exploiter la majorité des données brutes implique que les outils logiciels capables de réaliser des croisements, analyses et autres traitements de données en amont sont plus attractifs au regard des utilisateurs des plateformes de données (même si ces capacités en font des solutions difficiles à prendre en main et peu modulaires).

Le tableau que je dresse ici laisse entendre qu'il sera difficile aux plateformes de données territoriales de s'inscrire dans un écosystème d'outils interopérables et modulaires comme cela peut être le cas pour la cartographie. Si tant est que l'on s'en tienne à l'offre logicielle actuelle. Toutefois, compte tenu de la nature même des plateformes de données territoriales (variété des types de données, des profils d'utilisateurs et de leurs besoins) et de leurs stratégies, je suis persuadé que la voie à suivre est celle de la modularité et de l'interopérabilité.

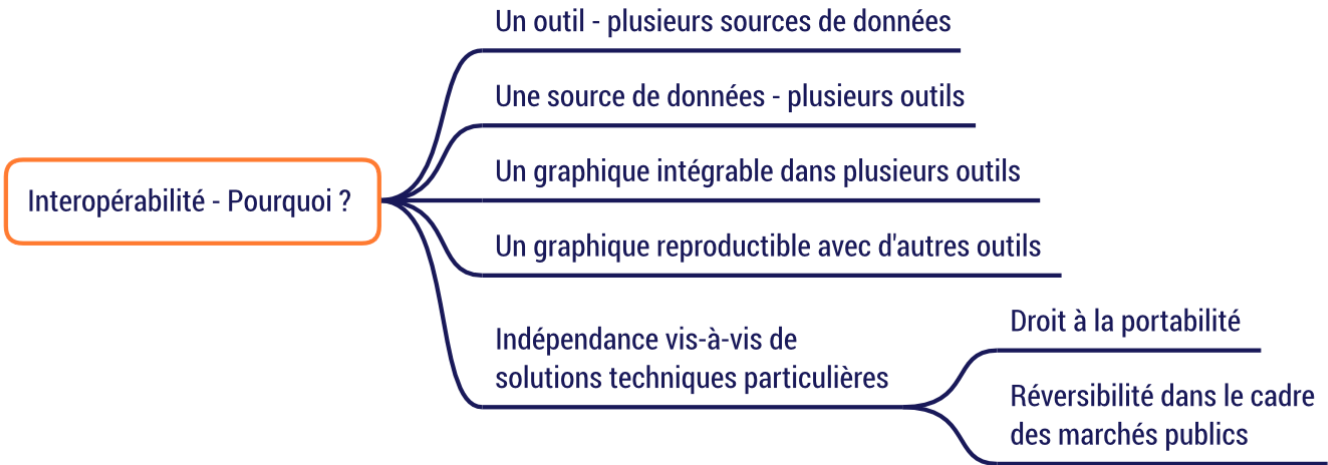
Même si le chemin de l'interopérabilité et de la modularité n'est pas le plus aisé, quelques signes récents montrent son intérêt. À titre d'exemples, je vous présente en quelques mots 2 initiatives menées par des plateformes régionales :

- La communauté Prodigé, après quelques expérimentations sur des voies totalement différentes, a entrepris de développer 2 modules distincts, l'un dédié à la réalisation de traitements et l'autre à la création de dataviz. Le premier permet de concevoir des indicateurs statistiques à partir de données brutes et le second consomme ces indicateurs pour présenter des graphiques. Ce choix de scinder en deux la composante dataviz est remarquable. En effet, il autorise d'autres outils, qu'ils soient internes ou externes, à ré-exploiter les services de calcul des indicateurs et à les présenter sous d'autres formes, voire à les rendre disponibles pour des usages hors dataviz. L'intérêt de la démarche est d'autant plus grand que les concepteurs de ces outils ont mis l'accent sur la réutilisabilité de ces services en implémentant le standard international [OGC API - Processes](#).
- DataGrandEst a développé un ensemble de web components dédiés à la dataviz : [DataGrandEst Dataviz Components](#). Un web component est un conteneur personnalisable et facilement réutilisable qui intègre du code et une interface graphique s'appuyant sur les standards du web (HTML, CSS, JavaScript). Avec les composants web de DataGrandEst l'accent est mis sur deux aspects : la création de dataviz simples et l'interopérabilité avec les sources de données classiques des plateformes de données (CSV, JSON, WFS et d'autres API). Ainsi, à l'aide de ces composants, la création de tableaux de bord simples externalisés et connectés à différentes sources de données est rendue accessible.



► aifigéo ◀

Construire
l'information géographique
de demain



Quelques motivations qui illustrent l'intérêt de l'interopérabilité.

Évidemment, l'interopérabilité et la modularité dans le domaine de la dataviz ne sont pas des fins en soi. Les acteurs impliqués dans le développement et l'intégration des solutions techniques adaptées aux plateformes de données devront approfondir ces réflexions en clarifiant les contextes d'usage et en identifiant les besoins fondamentaux autour desquels leurs communautés d'utilisateurs peuvent se reconnaître et se rejoindre. Dans cette optique, 2 axes de travail me semblent nécessaires :

- Développer les échanges entre plateformes de données à la fois pour partager les retours d'expérience et pour faire émerger des besoins cohérents en vue de mieux faire communiquer entre elles les solutions logicielles déployées sur leurs propres infrastructures ainsi qu'avec celles hébergées ailleurs ;
- Identifier des critères d'interopérabilité, de modularité et de réutilisabilité dans les cahiers des charges des plateformes.