



L'OPEN DATA AU SERVICE DE L'INVESTIGATION SCIENTIFIQUE PAR LES ÉLÈVES

Utiliser des données ouvertes afin de prolonger les investigations scientifiques



Dans le cadre du projet SteamCity, l'utilisation des données ouvertes représente une opportunité de prolonger les investigations scientifiques menées en classe en véritables projets de science citoyenne. Cette fiche pratique vise à accompagner les enseignants dans l'intégration de la démarche de l'open data à leurs pratiques pédagogiques, en montrant comment ces ressources permettent à la fois d'approfondir les démarches d'investigation et de valoriser le travail des élèves à travers le partage et la collaboration entre pairs.

Comprendre l'Open Data dans le contexte scolaire

Les données ouvertes, ou **Open Data**, désignent l'ensemble des données numériques accessibles librement, réutilisables sans restriction technique ou juridique, et partageables par tous.

Issus des pratiques scientifiques historiques, le mouvement de l'open data s'est structuré et démocratisé dans les années 2000 avec l'arrivée d'Internet et il a progressivement gagné d'autres domaines tel que l'administration publique et l'éducation.

Aujourd'hui, pour être considéré comme "ouvert", un jeu de données doit disposer de conditions d'utilisation qui respectent les principes établis par l'Open Knowledge Foundation comme la disponibilité et l'accès, la réutilisation et la redistribution, ainsi que la participation universelle. Ces conditions d'accès sont matérialisées les licences qui sont des documents juridiques représentant le contrat qui lie l'utilisateur avec le producteur d'une donnée.

Dans le contexte du projet SteamCity, les données ouvertes, en plus de représenter une source de connaissance extérieure, transforme la manière dont les élèves abordent l'investigation scientifique. Ils ne sont plus uniquement consommateurs d'informations mais deviennent aussi **producteurs de données qui enrichissent un patrimoine commun de connaissances sur leur territoire.**

Cette démarche s'inscrit dans l'ambition de SteamCity de développer une approche expérimentale pour les territoires apprenants. Les élèves peuvent contextualiser leurs observations dans un ensemble plus large de données, comparer leurs résultats avec ceux d'autres classes ou d'autres établissements, et contribuer activement à la compréhension des enjeux environnementaux et sociétaux de leur environnement proche. Cette approche répond aux recommandations de l'UNESCO sur l'éducation scientifique ouverte et aux objectifs de développement durable qui prônent une science plus inclusive et participative.

Recommandation de l'UNESCO sur une science ouverte

Ce document a pour but de fournir un cadre international pour les politiques et les pratiques de la science ouverte. Elle énonce une définition ainsi que des valeurs, des principes et des normes partagés à l'échelle internationale, et propose un ensemble d'actions favorisant une mise en place juste et équitable de la science ouverte pour tous.

https://unesdoc.unesco.org/ark:/48223/pf0000379949_fre



APPROFONDIR LA DÉMARCHE D'INVESTIGATION GRÂCE AUX DONNÉES OUVERTES

L'exploration initiale : construire sur l'existant

Avant de débiter l'expérimentation, l'exploration des données déjà disponibles constitue une étape fondamentale de la démarche scientifique. Les portails de données ouvertes, qu'ils soient locaux, nationaux ou européens, offrent des informations permettant aux élèves de contextualiser leur problématique et de mieux construire leurs questions de recherche. Les collectivités territoriales publient régulièrement des données sur la qualité de l'air, les transports, l'énergie ou encore la biodiversité.

Cette phase d'exploration permet aux élèves de formuler des hypothèses plus précises et de identifier les données manquantes qu'ils devront collecter. Elle développe leur esprit critique en les confrontant à la diversité des sources et à la nécessité d'évaluer la fiabilité et la pertinence des informations disponibles.

Les données disponibles doivent être adaptée à chaque situation et à chaque territoire. De manière générale, l'élève peut être encouragé à commencer par les sources de données les plus locales et à remonter progressivement vers des données plus larges.

Voici quelques **exemples multi-niveau de plateformes d'open data** pouvant être donné aux apprenants :

| Plateformes locales et régionales | Plateformes nationales |
|--|---|
| <ul style="list-style-type: none">Agglomération de La Rochelle : https://opendata.agglo-larochelle.fr/accueilGrand Lyon : https://data.grandlyon.comToulouse Métropole : https://data.toulouse-metropole.frParis : https://opendata.paris.frRégion Île-de-France : https://data.iledefrance.frRégion Sud : https://datasud.fr | <ul style="list-style-type: none">Portail national des données ouvertes : https://www.data.gouv.frGéorisques : https://www.georisques.gouv.frGéoportail IGN : https://www.geoportail.gouv.frINSEE : https://www.insee.frATMO France (qualité de l'air) : https://www.atmo-france.orgADES (eaux souterraines) : https://ades.eaufrance.frOpen Data Éducation : https://data.education.gouv.fr |

| Plateformes européennes | Plateformes internationales |
|--|--|
| <ul style="list-style-type: none">Portail européen de données : https://data.europa.euAgence Européenne de l'Environnement : https://www.eea.europa.eu/data-and-mapsQualité de l'air en Europe : https://www.eea.europa.eu/themes/air/air-qualitySystème d'information sur l'eau (WISE) : https://water.europa.euService d'observation du bruit : https://noise.eea.europa.euCopernicus Climate Data Store : https://climate.copernicus.eu/climate-data-storeCopernicus Atmosphere : https://atmosphere.copernicus.euCopernicus Land Monitoring : https://land.copernicus.eu | <ul style="list-style-type: none">OpenStreetMap : https://www.openstreetmap.orgWorld Bank Open Data : https://data.worldbank.orgOur World in Data : https://ourworldindata.orgGlobal Forest Watch : https://www.globalforestwatch.org |

L'enrichissement pendant l'expérimentation

Durant la phase de collecte de données, l'accès aux données ouvertes permet d'enrichir les observations des élèves en contextualisant leurs mesures dans un cadre plus large et en révélant des corrélations. Cette mise en perspective transforme une simple collecte de données en une véritable investigation scientifique multidimensionnelle.

Voici, plusieurs **exemples de protocoles** qui pourraient être enrichis grâce aux données ouvertes :

- Sur **l'étude de la qualité de l'air** autour de l'établissement scolaire, projet particulièrement pertinent pour travailler autour du bien vivre sur son territoire. Les mesures de particules fines effectuées par les élèves avec des capteurs peuvent être immédiatement confrontées aux données horaires des stations des agences de qualité de l'air les plus proches, permettant de valider le bon fonctionnement des instruments et d'identifier d'éventuels phénomènes hyperlocaux. Le croisement avec les données de comptage routier disponibles sur les portails municipaux révèle souvent des corrélations entre les pics de pollution et les heures de pointe du trafic. L'intégration des données météorologiques, notamment la direction et la force du vent issues des stations Météo nationale, permet de comprendre la dispersion des polluants et d'expliquer certaines variations apparemment aléatoires.
- Un projet **d'étude de la biodiversité urbaine** illustre différemment cette synergie entre données collectées et données ouvertes. Pendant que les élèves inventorient les espaces propices aux pollinisateurs dans différents espaces verts de leur commune, ils peuvent enrichir leurs observations avec les données du programme CORINE Land Cover (<https://land.copernicus.eu/en/products/corine-land-cover>) qui caractérise l'occupation des sols à l'échelle européenne. Cette confrontation permet de comprendre comment la fragmentation des habitats influence la présence de certaines espèces. Les données historiques d'urbanisation, souvent disponibles via les archives cadastrales numérisées, révèlent l'évolution du paysage et expliquent la répartition actuelle de la biodiversité.
- **L'analyse des îlots de chaleur urbains** constitue un autre exemple particulièrement démonstratif de cette approche intégrée. Les relevés de température effectués par les élèves dans différents quartiers prennent une nouvelle dimension lorsqu'ils sont croisés avec les données satellites thermiques du programme Copernicus qui offrent une vision large du phénomène. Les bases de données urbaines sur la densité du bâti, la hauteur des immeubles et la surface de végétalisation, généralement disponibles en open data, permettent d'établir des corrélations précises entre morphologie urbaine et température. Les données de consommation énergétique agrégées par quartier, publiées par de nombreuses métropoles dans le cadre de leurs plans climat, révèlent le lien entre îlots de chaleur et demande énergétique pour la climatisation.
- Un projet sur la **mobilité durable** offre encore d'autres perspectives d'enrichissement. Les comptages de cyclistes effectués par les élèves sur des axes stratégiques peuvent être comparés aux données des compteurs automatiques permanents installés par de nombreuses villes et dont les résultats sont publiés en temps réel. Les données d'accidentologie routière, disponibles sur data.gouv.fr avec une géolocalisation précise, permettent d'identifier les points noirs de sécurité et de comprendre les réticences à l'usage du vélo. Les temps de parcours en transports en commun, accessibles via les API des opérateurs de transport, offrent des éléments de comparaison objectifs entre différents modes de déplacement. Les données de qualité de l'air le long des axes routiers permettent de quantifier l'exposition des cyclistes et des piétons à la pollution.

Cette approche développe chez les élèves une compréhension systémique des phénomènes étudiés. Ils découvrent que chaque problématique environnementale et/ou territoriale s'inscrit dans un réseau complexe d'interactions où les facteurs naturels, anthropiques, historiques et socio-économiques s'entremêlent. La capacité à mobiliser et croiser différentes sources de données devient alors une compétence essentielle pour appréhender cette complexité. Les élèves apprennent ainsi que l'investigation scientifique ne se limite pas à l'observation directe mais s'enrichit de la capacité à naviguer dans l'écosystème des données ouvertes pour construire une compréhension systémique des phénomènes étudiés.

L'analyse approfondie des résultats

La phase d'analyse représente un moment privilégié pour développer l'esprit critique des élèves et leur compréhension de la démarche scientifique. L'exploitation des données ouvertes offre ici une opportunité d'aborder la question fondamentale de la qualité des données et de la validité des résultats scientifiques.

Les résultats obtenus peuvent être validés par comparaison avec des données de référence, ce qui permet aux élèves de questionner leur méthodologie et d'affiner leur compréhension du protocole scientifique. L'élargissement de la perspective par le croisement de différentes sources de données révèle des corrélations inattendues et stimule la réflexion sur les causalités. La complexité territoriale des phénomènes étudiés devient accessible grâce à la multiplicité des angles d'analyse possibles.

Durant cette étape, il est important pour l'enseignant de questionner le travail d'investigation pour que l'élève en comprenne la portée mais aussi puisse en saisir les biais ainsi que les limites. Pour cela, il peut aborder les aspects suivants :

- **La comparaison avec les données de référence** : Non pas comme un jugement de valeur, mais comme une opportunité d'analyser les sources de variation (précision des instruments, conditions de mesure, protocoles de calibration)
- **L'importance des métadonnées** : Expliquer comment la documentation exhaustive des conditions de production des données est essentielle pour garantir leur reproductibilité et leur comparabilité.
- **L'apprentissage statistique concret** : Présenter l'utilisation des notions de moyennes, écarts-types et intervalles de confiance sur les données d'expérimentation de pour comprendre l'incertitude inhérente à toute mesure.
- **La représentativité spatiale et temporelle** : Expliquer comment une mesure ponctuelle ne peut prétendre représenter l'ensemble d'un territoire, et bien insister sur l'importance des réseaux de mesure.
- **L'interopérabilité et la standardisation** : Présenter cet aspect de la donnée non comme une contrainte mais comme une condition pour la collaboration scientifique entre pairs.
- **La valorisation de la transparence scientifique** : Insister sur le fait que reconnaître les limites de ses données est une marque de maturité scientifique et non un aveu de faiblesse.

L'enseignant joue ici un rôle déterminant en guidant cette analyse critique sans décourager les élèves. Chaque investigation, aussi modeste soit-elle, contribue à l'édifice de la connaissance pour autant qu'elle respecte ces principes de qualité et de transparence. Cette perspective encourage les élèves à améliorer continuellement la méthodologie, à affiner les protocoles, et à considérer chaque expérience comme une étape dans un processus de recherche de la connaissance plutôt que comme une fin en soi.



RENDRE LE TRAVAIL D'ENQUÊTE VISIBLE ET PARTAGEABLE

La structuration des données pour le partage

La valorisation du travail d'investigation passe d'abord par une structuration rigoureuse des données collectées. Cette exigence de rigueur constitue en elle-même un apprentissage fondamental de la démarche scientifique. Les données doivent être organisées selon un format standardisé incluant systématiquement la date, l'heure, le lieu précis de la mesure, la valeur mesurée avec son unité, le type de capteur utilisé, l'identification de la classe et tout commentaire pertinent sur les conditions de mesure.

L'utilisation de formats ouverts comme le CSV ou le JSON garantit l'interopérabilité des données et leur réutilisation par d'autres acteurs. La documentation de la méthodologie employée, à travers une fiche explicative détaillée, assure la reproductibilité de l'expérience et la compréhension du contexte de collecte. L'attribution d'une licence Creative Commons explicite les conditions de réutilisation et sensibilise les élèves aux enjeux de la propriété intellectuelle dans le domaine scientifique.

Les outils de visualisation et de diffusion

La transformation des données brutes en visualisations compréhensibles constitue une étape de la communication scientifique. Des outils comme Datawrapper (<https://www.datawrapper.de/>) permettent de créer facilement des graphiques et des cartes interactives adaptés au niveau des élèves. Pour des projets plus ambitieux, uMap (<https://umap.openstreetmap.fr/fr/>) offre la possibilité de créer des cartes personnalisées s'appuyant sur OpenStreetMap, tandis que Grafana (<https://grafana.com/>) permet de construire des tableaux de bord dynamiques pour un suivi en temps réel des mesures.

Le choix de l'outil dépend du niveau des élèves et de la complexité du projet, mais l'objectif reste constant : rendre les données accessibles et compréhensibles pour différents publics, des pairs aux décideurs locaux.

La création de valeur par la collaboration

Le partage des données ouvre des perspectives de collaboration particulièrement enrichissantes. La comparaison synchrone, où plusieurs classes effectuent simultanément des mesures dans différents établissements, quartiers ou villes, permet de cartographier en temps réel les variations territoriales d'un phénomène. Cette approche développe le sentiment d'appartenance à une communauté scientifique élargie et donne du sens aux efforts de standardisation des protocoles.

La dimension internationale du projet prend ici tout son sens. L'échange de données avec des partenaires européennes permet d'analyser comparativement l'impact de politiques environnementales différentes et de comprendre comment les contextes géographiques, culturels et réglementaires influencent les problématiques territoriales.

La contribution au patrimoine de données citoyennes représente l'aboutissement de cette démarche. La publication des données sur les portails open data locaux permet leur utilisation par d'autres acteurs du territoire : associations environnementales, chercheurs universitaires, services municipaux. Les élèves prennent ainsi conscience de leur capacité à contribuer activement à la connaissance collective et à influencer à leur échelle les décisions territoriales.



MISE EN ŒUVRE PRATIQUE : L'EXEMPLE D'UNE CARTOGRAPHIE SONORE

Pour illustrer concrètement cette approche, considérons un projet de cartographie sonore du quartier de l'établissement. Ce projet se déroule en quatre phases distinctes sur environ six séances.

La phase de **préparation** commence par l'exploration des cartes de bruit existantes, généralement disponibles sur les portails open data municipaux conformément à la directive européenne sur le bruit environnemental. Cette analyse permet d'identifier les zones d'intérêt et les points de comparaison pertinents. Les élèves définissent ensuite un protocole de mesure rigoureux, précisant les horaires, les durées et les conditions de mesure.

La phase de **collecte** mobilise les élèves sur le terrain avec des sonomètres ou des applications mobiles calibrées. L'enregistrement structuré des données s'accompagne de prises de vue photographiques et d'observations qualitatives sur l'environnement sonore. Cette phase développe les compétences d'observation et la rigueur expérimentale.

Le **traitement des données** implique leur importation dans un tableur pour analyse statistique, la création d'une carte interactive permettant de visualiser les niveaux sonores, et la comparaison avec les normes réglementaires en vigueur. Cette phase mobilise les compétences numériques et mathématiques des élèves.

La phase de **partage et de valorisation** voit les résultats publiés sur la plateforme SteamCity ou Vittamap, présentés aux autres classes du réseau, et potentiellement soumis au conseil municipal sous forme de propositions d'amélioration de l'environnement sonore. Cette restitution développe les compétences de communication et l'engagement citoyen.



LES COMPÉTENCES DÉVELOPPÉES

L'intégration de l'open data dans les investigations scientifiques développe un ensemble cohérent de compétences essentielles pour former les citoyens de demain.

Sur le plan scientifique, les élèves acquièrent une rigueur méthodologique dans la collecte et le traitement des données. Ils développent leur capacité d'analyse critique des sources d'information et apprennent à interpréter des ensembles de données complexes en identifiant les tendances, les corrélations et les anomalies. La confrontation avec des données réelles les sensibilise aux incertitudes et aux limites inhérentes à toute mesure scientifique.

Les compétences numériques mobilisées vont bien au-delà de la simple utilisation d'outils. Les élèves apprennent à manipuler différents formats de fichiers, à transformer des données brutes en visualisations pertinentes, et à publier de manière responsable sur internet en respectant les principes de l'open data et les contraintes juridiques associées.

Les compétences citoyennes se construisent à travers la compréhension concrète des enjeux territoriaux. La participation active à des projets de science citoyenne développe le sentiment d'efficacité personnelle et collective face aux défis environnementaux et sociétaux. L'engagement dans la vie locale, par la production de données utiles à la communauté, renforce le lien entre l'école et son territoire.



POINTS DE VIGILANCE ET RECOMMANDATIONS

La mise en œuvre de projets impliquant l'open data nécessite une attention particulière à certains aspects juridiques et pédagogiques.

Sur le plan juridique, le respect du Règlement Général sur la Protection des Données impose de veiller à ne jamais collecter ni publier de données personnelles. Les élèves doivent être sensibilisés à ces enjeux dès la conception du protocole expérimental. La vérification des licences des données utilisées garantit le respect de la propriété intellectuelle et enseigne les bonnes pratiques du partage de connaissances. L'obtention des autorisations nécessaires pour les mesures dans l'espace public implique un dialogue avec les autorités locales qui peut lui-même être formateur.

Sur le plan pédagogique, l'adaptation de la complexité au niveau des élèves reste primordiale. Il convient de privilégier la compréhension de quelques aspects plutôt que le traitement superficiel d'une grande quantité de données. La valorisation de la démarche scientifique, indépendamment de la spectacularité des résultats, encourage la persévérance et l'honnêteté intellectuelle.



CONCLUSION

L'intégration de l'open data dans les investigations scientifiques scolaires représente bien plus qu'un simple enrichissement méthodologique. Elle transforme fondamentalement la relation des élèves à la connaissance scientifique et à leur territoire. En devenant producteurs de données ouvertes, les élèves contribuent activement à la construction d'un bien commun informationnel. En exploitant les données existantes, ils inscrivent leurs observations dans un contexte élargi qui donne du sens à leur travail. En partageant leurs résultats, ils participent à une dynamique collaborative qui dépasse les murs de la classe et les frontières nationales.

Le projet SteamCity offre le cadre idéal pour cette transformation pédagogique. Il fournit les outils, les ressources et le réseau nécessaires pour que chaque investigation scolaire devienne une contribution significative à la compréhension des enjeux territoriaux. L'enseignant, dans ce contexte, devient un facilitateur qui accompagne les élèves dans leur transformation en citoyens-scientifiques capables de comprendre, produire et partager des données pour agir sur leur territoire.

Cette approche prépare les élèves aux défis du XXI^e siècle en développant simultanément leur culture scientifique, leurs compétences numériques et leur conscience citoyenne. Elle fait de l'école un véritable laboratoire d'innovation territoriale où se construisent les solutions de demain.