

L'OpenData : découverte d'un concept

Conditions	Travail en groupe Durée : 2h	Moyens	<ul style="list-style-type: none"> • Poste informatique sous Windows • Connexion internet
Prérequis	aucun	Niveau	Classe de Terminale S spécialité ISN
Compétences	<ul style="list-style-type: none"> • C1.1 Justifier dans une situation donnée, un codage numérique ou l'usage d'un format approprié... • C1.2 Détailler le déroulement d'une communication numérique... • C4.1 Documenter un projet numérique pour en permettre la communication... • C5.1 Avoir conscience de l'impact du numérique dans la société notamment de l'importance des licences et du droit. • C5.2 Mesurer les limites et les conséquences des lois régissant les échanges numériques. 		
Éléments du programme	<ul style="list-style-type: none"> • Représentation de l'information <ul style="list-style-type: none"> ○ Formats ○ Structuration et organisation de l'information ○ Non-rivalité de l'information • Architectures matérielles <ul style="list-style-type: none"> ○ Réseaux 		

Partout dans le monde, les états et les collectivités locales s'engagent dans le mouvement OpenData. Ce phénomène est fortement lié à l'avènement du web2.0 et nous laisse entrevoir notre futur proche à travers la notion de « smartcity » connue aussi sous l'appellation « ville2.0 ».

En quoi consiste l'OpenData ? Qu'est-ce qu'une donnée ouverte ? Y-a-t-il des règles, des lois qui régissent l'OpenData ? Comment utiliser les données ? Comment y accéder ?

Table des matières :

1. Introduction à l'Opendata :	2
2. L'Opendata et son vocabulaire :	2
3. Plateforme NOD : Nantes Ouvertures des Données	4
Quelles données ?	4
Fonctionnement de l'API.....	5
Accès aux données sous quelle licence ?	6
Création d'un compte sur la plateforme NOD	7

1. Introduction à l'OpenData :

L'OpenData c'est rendre accessible au citoyen les données publiques, dans le cadre des grandes métropoles, des gouvernements, et des organismes publics de recherche. L'objectif est de développer des applications innovantes, d'améliorer les services et de créer de « la croissance » économique. C'est aussi une façon d'assurer la transparence publique.

Visualiser le spot de présentation de la démarche d'ouverture des données numériques publiques (initié par LiberTIC, soutenu par Nantes Métropole et réalisé par A2B Production en licence Art Libre) : « [L'Open Data à la loupe](#) »

Résumer le concept d'OpenData en répondant aux questions suivantes :

1. Lister les quatre catégories de données mise à disposition sur les plateformes OpenData :
Les données statistiques, cartographiques, les horaires, les données économiques
2. Les données publiques sont rendues accessibles contrairement aux données privées. En quoi une donnée est dite privée ?
Données privées = données sensibles et à caractère personnel
3. A qui s'adresse l'OpenData ?
Chercheurs, journalistes, étudiants, citoyens
4. Quelles sont les données recoupées par l'application Opendata destinée à « faciliter l'accès aux transports aux personnes à mobilité réduite » ?
Horaires de bus, géolocalisation, hauteur des trottoirs
5. Sur quel appareil l'application pour « faciliter l'accès aux transports aux personnes à mobilité réduite » est-elle consultable ?
smartphone

2. L'OpenData et son vocabulaire :

L'OpenData, comme tous les domaines techniques, utilise un vocabulaire spécifique dont il faut connaître les usages afin d'en comprendre les origines et surtout les perspectives.

A l'aide d'une recherche sur le web, définir les termes suivants :

6. API
API (Application Programming Interface) désigne une interface de programmation fournie par un programme informatique pour permettre à d'autres programmeurs d'en utiliser certaines fonctions ou données. A travers le développement d'API, il s'agit pour un programmeur de développer une fois pour toute une application contenant tous les outils de base nécessaires, et de la mettre à disposition de la communauté des développeurs pour leur faciliter le travail.
7. Crowdsourcing
Désigne le fait d'utiliser l'intelligence, la créativité, le savoir-faire d'un grand nombre de personnes (en l'occurrence des internautes) pour enrichir, alimenter un service, une application, un site web, la conception d'un produit, etc.
8. Data journalism

Désigne une nouvelle façon de faire de l'investigation journalistique en se basant sur l'analyse de données et (souvent) la représentation visuelle. Le journaliste utilise des bases de données comme sources et en déduit des connaissances, des corrélations ou des intuitions qui ne seraient pas accessibles par les méthodes traditionnelles de l'enquêtes journalistique.

9. Linked Data

Désigne une approche du web poussée par les tenants du "web sémantique" qui encourage à décrire toutes les données d'une manière lisible par des ordinateurs, et à les lier entre elles en décrivant leurs relations, ou encore, en facilitant leur mise en relation : ainsi, deux informations produites par des acteurs différents, mais correctement décrites comme faisant référence à un même territoire, pourront être "liées" même si leurs producteurs n'y auraient jamais pensé.

10. Interopérabilité

Désigne la capacité que possède un produit ou un système, dont les interfaces sont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre

11. Métadonnée

Désigne un ensemble structuré d'informations servant à définir ou décrire une ressource (donnée, document...) quel que soit son support (papier ou électronique). S'agissant d'un livre par exemple, le titre, le nom d'auteur, l'éditeur, l'année de publication en constituent des métadonnées. Les métadonnées sont donc "des données sur les données", qui permettent à un individu ou un ordinateur d'en comprendre le sens et l'organisation. Elles sont à la base des techniques du web sémantique.

12. Web 2.0

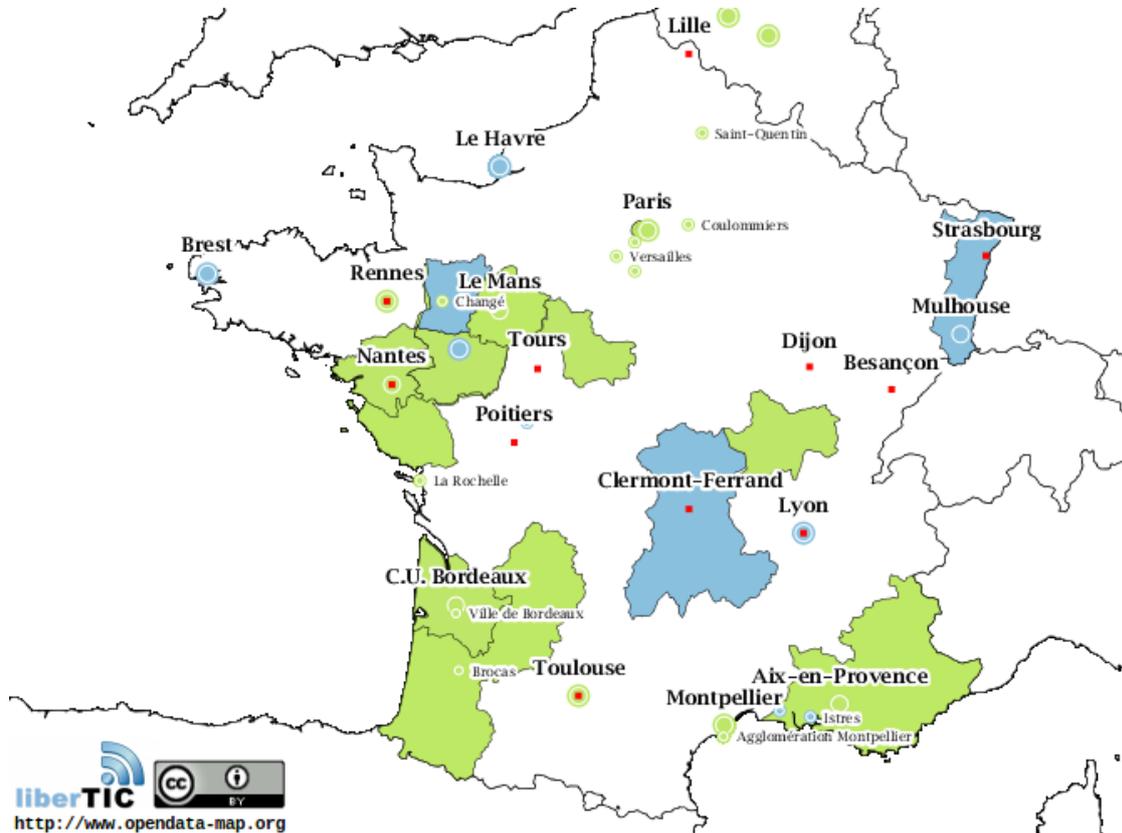
Il n'existe pas de définition stable. L'expression a été proposée en 2004 par l'éditeur Tim O'Reilly en 2004 pour décrire une certaine évolution du web fondée sur quatre piliers :

- Le recentrage de l'information et des services autour de la personne ;
- Une facilitation considérable de l'intervention des utilisateurs dans la production et l'échange de textes, images et vidéos ;
- Des interfaces standards permettant de partager et d'agencer de diverses manières informations et services ;
- Et une expérience plus "sociale" du web par laquelle l'innovation et la valeur proviennent, dans une large mesure, de l'interaction entre les utilisateurs.

Les services et plateformes emblématiques du web2.0 sont Wikipedia, eBay et Craigslist, Youtube, Facebook et Twitter.

3. Plateforme NOD : Nantes Ouvertures des Données

Nantes Métropole et la Ville de Nantes ont ouvert leurs données en février 2011 avec le site <http://data.nantes.fr>. Ils ont suivi le mouvement engagé par la ville de Rennes et la ville de Paris. Aujourd'hui de nombreuses collectivités se lance dans le projet d'ouverture de leurs données. Voici la [carte](#) de l'opendata en France.



Légende: ■ Déjà ouvert ■ En cours ■ Mouvement citoyen

Quelles données ?

Connectez-vous à la plateforme nantaise d'accès aux données et rendez-vous dans la rubrique Données puis dans la sous-rubrique Statistiques des données.

NANTES OUVERTURE DES DONNÉES

ouverture des données publiques

Accueil
Données
Actualités
Démarche
Licence
Appel à projets
Forum
Mon compte

Choix des formats
Fonctionnement de l'API
Statistiques des données

13. Relever les 4 principales catégories de données proposées.

Santé/Sociale (19%), Citoyenneté/Institutions (18%), Mobilité (17%), Culture/Tourisme (11%)

14. Relever les 4 principaux formats de données proposées.

XLS, CSV, JSON, XML

15. Parmi les jeux de données proposées, lesquelles sont mises à jour en temps réel, quotidiennement et mensuellement.

Données temps réel :

- API temps réel de la TAN
- Disponibilité dans les parkings publics de Nantes Métropole
- Fluidité des axes routiers de Nantes Métropole
- Info-traffic TAN
- Temps de parcours sur les itinéraires de Nantes Métropole.

Données quotidiennes : Fiches descriptives d'ensembles et sous-ensembles d'articles d'archives des archives régionales des Pays de la Loire

Données mensuelles :

- ActiCarte en Pays de la Loire
- Destinéo en Pays de la Loire
- Encours de crédits bancaires en Pays de la Loire
- Fréquentation des bacs de Loire en Loire-Atlantique
- Impôt sur les sociétés en Pays de la Loire
- Liste des tarifs applicables dans les équipements culturels de Nantes
- Prime Régionale à la Création d'Entreprise Jeunes et Solidaires (PRCE-Jeunes et Solidaires) en Pays de la Loire
- Taxe Intérieure de Consommation sur les Produits Énergétiques (TIPCE) en Pays de la Loire
- Taxe sur la Valeur Ajoutée (TVA) en Pays de la Loire

Fonctionnement de l'API

Rendez-vous dans la rubrique Données puis dans la sous-rubrique Fonctionnement de l'API.

Deux types d'API sont mis à disposition : la première, l'API v2, pour interroger l'ensemble des jeux de données et la seconde, l'API v1, pour récupérer les données actualisées en temps réel.

16. Relever le modèle de construction d'un appel sur l'API v2

`http://data.<nom_collectivité>.fr/api/publication/<nom_jeu>/<nom_table>/content?filter={<filtre>}&key=<clé>&format=<nom_format>&limit=<nombre_max_lignes>&offset=<valeur_offset>`

17. Indiquer le format par défaut des réponses de l'API v2

Par défaut, le format utilisé est JSON

18. Relever le modèle de construction d'un appel sur l'API v1

`http://data.nantes.fr/api/<cmd>/<version>/<key>`

19. Indiquer le format par défaut des réponses de l'API v1

L'API fournit des réponses au format XML

Accès aux données sous quelle licence ?

Le cadre légal propre à l'exploitation des données publiques est défini par la loi du 17 juillet 1978

Elle reconnaît à toute personne un droit d'accès et de réutilisation des informations publiques.

L'organisme en charge de la veille au respect de la loi sur l'accès aux documents administratifs est la CADA Commission d'Accès aux Documents Administratifs.

20. Indiquer le type de licence utilisée pour l'Open Data de Nantes.

Une licence dite «libre» : la licence Open Data Base License (ODbL)

Il existe cependant, en France, des données publiques payantes répertoriées dans le document disponible en suivant ce [lien](#).

21. Rechercher dans ce document, le prix de la licence pour l'accès aux prix du carburant dans les stations-services.

145 K€ en 2011 prix de la licence pour accéder aux données sur le prix des carburants pratiqués dans les stations-services distribuant au minimum 500 m3 de carburants par an

22. Rechercher la différence entre les licences libres et celles gratuites.

Source : http://fr.wikipedia.org/wiki/Licence_libre

Un des aspects souvent mal compris du principe de licence libre est que celui-ci ne traite pas de la valeur marchande de la diffusion des œuvres. Une œuvre sous licence libre n'est pas nécessairement disponible gratuitement, pas plus qu'une œuvre disponible gratuitement n'est nécessairement libre (cette confusion est entretenue par le double-sens du mot anglais *free*, c'est pourquoi le terme «libre» est employé même dans des États non-francophones).

Le principe de licence libre n'interdit pas en effet de faire payer l'accès à l'œuvre, il garantit juste des libertés sur l'œuvre une fois celle-ci obtenue. Cela est d'autant moins bien compris que dans les faits, la majorité des œuvres sous licences libres sont disponibles gratuitement. D'autant que si une œuvre sous licence libre n'est au départ disponible que contre paiement, dès la première diffusion l'œuvre peut être rediffusée gratuitement en toute légalité.

23. Rechercher sur le web le principe de la licence Creative Commun.

Les **licences Creative Commons** constituent un ensemble de licences régissant les conditions de réutilisation et/ou de distribution d'œuvres (notamment d'œuvres multimédias diffusées sur Internet).

Le système se base sur plusieurs paramètres binaires :

- commercial / non commercial (NC) ;
- modifiable / non modifiable (ND) ;
- créations dérivées à partager selon la même licence (SA) / licence au choix du créateur final.

Aussi avons-nous un arbre de possibilités contenant $2 \times 2 \times 2 = 8$ combinaisons possibles.

Mais les deux combinaisons incluant à la fois les conditions ND (non modifiable) et SA (créations dérivées à partager selon la même licence) ne sont pas valides, car ces deux conditions s'excluent. Il reste ainsi six solutions : CC-BY, CC-BY-SA, CC-BY-ND, CC-BY-NC-SA et CC-BY-NC-ND

Création d'un compte sur la plateforme NOD

Pour pouvoir interroger la base de données, il faut s'inscrire sur le site afin créer un compte. Puis obtenir une clé d'application en remplissant un petit formulaire.



Avec la clé d'application créée, nous pourrons réaliser un petit programme destiné à interroger l'API.