

S-14 Données ouvertes et Big Data

Basma Makhlouf Shabou

Plan

- Données ouvertes
 - Terminologie
 - Principes
 - Enjeux et considérations archivistiques
- Big Data
 - Terminologie
 - Caractéristiques
 - Enjeux et considérations archivistiques
- Discussion: rôle de l'archivistique
- Approches d'enseignement

Données ouvertes - terminologie

- **Données ouvertes:** «Données brutes non nominatives et libres de droits, produites ou recueillies par un organisme public ou privé, qui sont accessibles aux citoyens par Internet» (Office québécois de la langue française 2013).

Les données ouvertes en vidéo

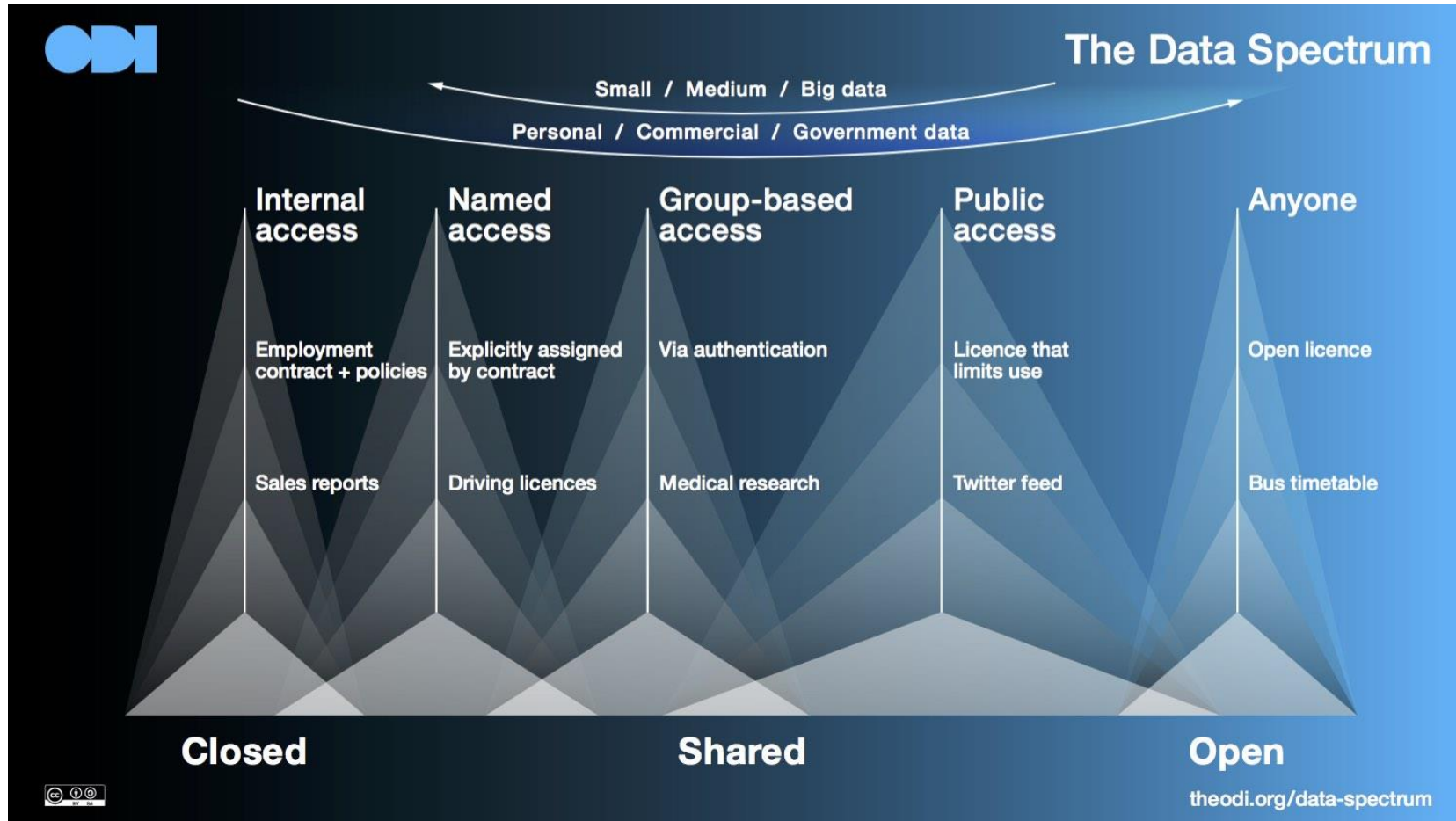


L' Open Data expliqué en emojis



opendata.swiss - le portail des données ouvertes de l'administration publique suisse (OGD)

The Data Spectrum



Données ouvertes – principes (Open Data Charter [s.d.]



Ouvertes par défaut



Diffusées en temps opportun et exhaustives



Accessibles et utilisables



Comparables et interopérables



Servant à l'amélioration de la gouvernance et de la participation citoyenne



Favorisant le développement inclusif et l'innovation

Données ouvertes - Enjeux et considérations archivistiques

■ Conservation permanente

- Dispositifs et garanties de l'accès et de l'exploitabilité durables des données ouvertes

■ Évaluation

- Identifiabilité des données qui ne peuvent pas être partagées (données sensibles)

■ Protection et droits d'accès

- Données sensibles 

Big Data - terminologies

- **Big Data:** «désigne à la fois la production de données massives et le développement de technologies capables de les traiter afin d'en extraire des corrélations ou du sens» (Grinbaum 2017, p. 2).

Le Big Data en vidéo



Caractéristiques du Big Data (Imbert 2016)

- 3 V
 - **V**olume : peta, exa, zettaoctet...
 - **V**élocité : fréquence à laquelle les données sont générées...
 - **V**ariété : données brutes, structurées, texte, image...
- + 2 autres:
 - **V**éracité: précision, pertinence → fiabilité
 - **V**aleur: génération d'une valeur pour l'institution

Big Data - Enjeux et considérations archivistiques

■ Contrôle et gouvernance

- Mise en place de stratégies par les gouvernements, par ex. la Maison Blanche (NSTC, 2016: <https://www.nitrd.gov/pubs/bigdatardstrategicplan.pdf>)

■ Conservation permanente

- Dispositifs et garanties de l'accès et de l'exploitabilité durables des données massives

■ Évaluation et élimination

- Identifiabilité des données à valeur permanente / données éphémères

■ Protection

- Données récoltées à travers différentes applications (Internet des Objets)
- Recoupage de plusieurs sets données anonymisées peut amener à une perte d'anonymisation (PFPDT [s.d.])

Discussion: rôle de l'archivistique

Selon vous, quel rôle les archivistes et gestionnaires de documents d'activité peuvent jouer face aux enjeux des données ouvertes et au Big Data?

Approches d'enseignement

Selon vous, quelles approches (disciplinaires, pédagogiques et techniques) conviendraient à l'enseignement de la gestion des données massives et données ouvertes?

Merci!



Questions?



“Study School for Archival Educators” copyright International Council on Archives, 2019, is licensed under Creative Commons License Attribution-Noncommercial 4.0.