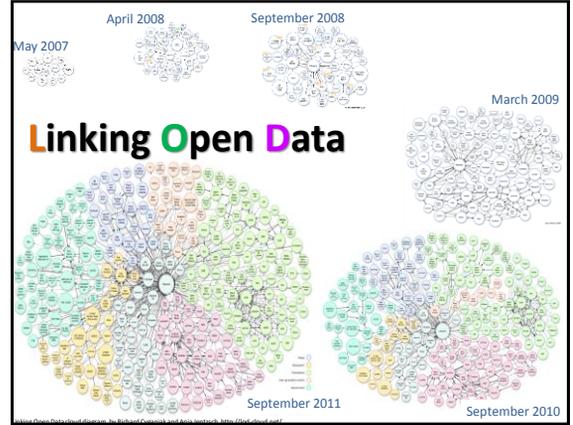
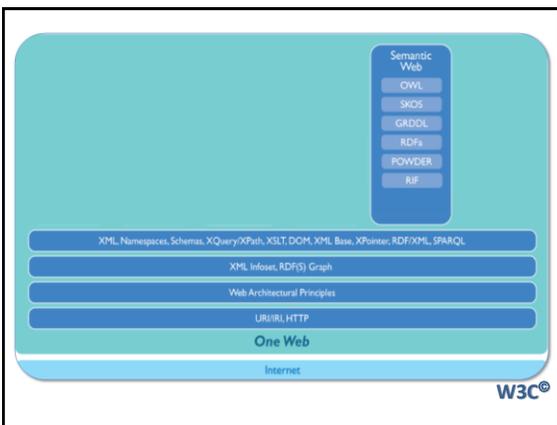
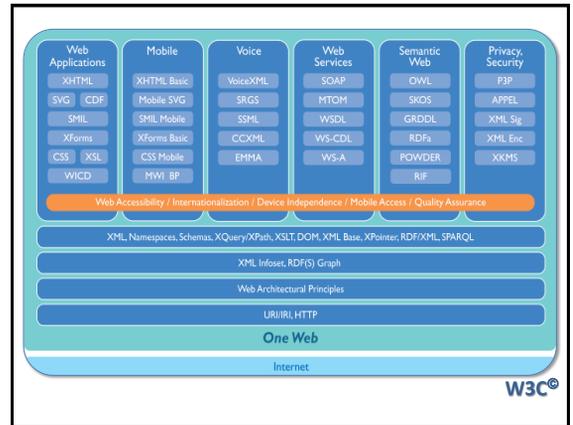


**DATALIFT**

libérez! les données! libérez! les données! libérez! les données! libérez!  
 les données! libérez! les données! libérez! les données! libérez! les  
 données! libérez! les données! libérez! les données! libérez! les données!  
 libérez! les données! libérez! les données!



**e.g. DBpedia**



**principes**

- utiliser le format **RDF**
- utiliser des **URI** pour nommer
- utiliser des URI **HTTP** (URL) pour que l'on puisse les suivre
- lorsqu'un URI est **suivi** donner des informations (HTML, RDF)
- inclure dans ces informations des **liens** vers d'autres données



**le cas de la France** mouvement d'ouverture des données mais...

- redondance des identifiants
- hétérogénéité des formats
- diversité des vocabulaires
- hétérogénéité des accès

Carte de France des initiatives d'ouverture de données publiques (Mars 2011 - Mars à Juin de la semaine 2012) EMI FRANÇAIS DATA.GOV.FR

**enjeu de data lift**

passer de la donnée brute publiée à la donnée sémantique interconnectée

**faciliter l'ouverture des données**

assister la publication sur le web et aux standards du web

**data - lift**

simplifier le processus de lifting / d'élévation de données  
sélection, formatage, schémas, liage, droits, licences, etc.

- principes de conversion pour les identifiants
- outils de conversion vers RDF

**sélection des données & conversion**



- catalogue, méthodes et indicateurs pour la sélection d'ontologies / schémas
- conversion RDF 2 RDF: restructuration, nettoyage et homogénéisation

**schémas, sémantique & conversion**



- architecture web standardisée open-source
- composants d'aide à la publication
- configuration de publication
- licences et droits d'accès aux données



**publication adaptée des données**



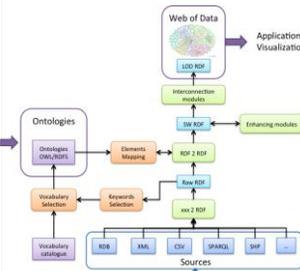
- composants d'alignement de ressources
- outils statistiques, calculs de clefs, etc.
- paramétrage et automatisation





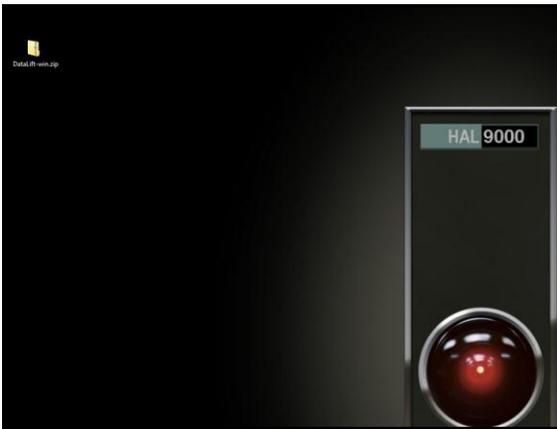
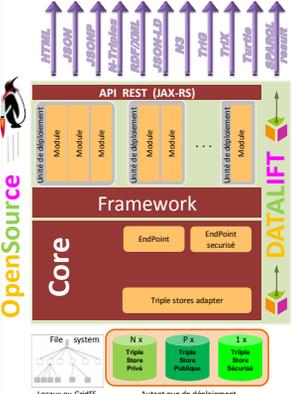
**interconnexion, liage de données**

### publication processus



### démo

- installation en un click
- import de données trouvées en ligne
- transformation en RDF
- publication sur le web
- interrogation directe

**architecture**

- standards web
- open source
- modulaire...

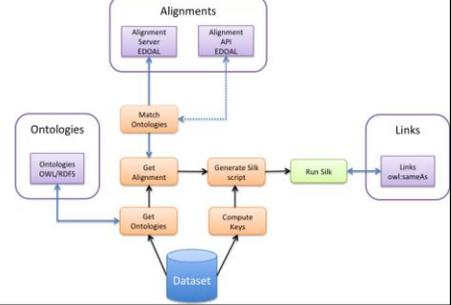
### module LOV (Linked Open Vocabularies)

- catalogue des vocabulaires réutilisables
- sélection par des experts de plus de 260 vocabulaires (20 000 termes) certifiés Datalift
- accessible dans la plateforme: réutilisation & extension
- partagé sur le Web <http://labs.mondeca.com/dataset/lov/>

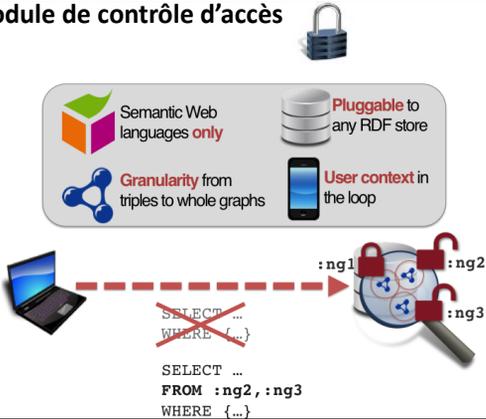


### module interconnexion de données

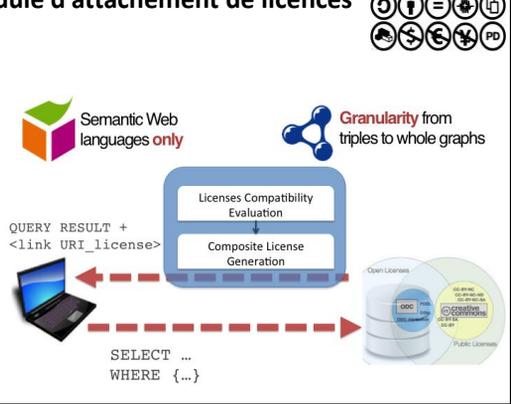
- automatiser l'interconnexion
- utilisation des alignements entre ontologies
- calcul des clefs sur les jeux de données pour réduire le nombre de comparaisons
- génération automatique de spécifications pour l'outil Silk



### module de contrôle d'accès



### module d'attachement de licences



### phase 2: mise en ligne de jeux de données validation et application de la plate-forme à des données réelles



datalift.org

