

Innover par et
pour la donnée

Logilab

@ Bibcamp'2015

nicolas.chauvat@logilab.fr

Fondateur et PDG de Logilab depuis l'an 2000

Vous n'êtes pas du métier ?

ma culture =

recherche en informatique

+ logiciel libre

+ Web

+ agilité

Comment êtes-vous arrivé là ?

- CubicWeb débuté en 2001
- Linked Open Data pour des musées en 2006
- Rencontré BnF à une conference en 2007
- Gagné marché en 2010 puis 2012
- Depuis de nombreux projets intéressants

Vous vendez quoi ?

- conseil et formation
(urbanisation des données, web sémantique, agilité, etc.)
- publication web sémantique
(culture = opencat + fevis + biblissima + musées)
- traitement des données
(nettoyage, enrichissement, alignement, etc.)

data.bnf

le Web



LICENCE OUVERTE
OPEN LICENCE

Start small, evolve step by step

Schema.org

SPARQLendpoint



| 2011

| 2012

| 2013

| 2014

**5 000
authors**

**Music works
Subject pages
(RAMEAU)**

**200 000
authors**

**Performances
Geographic
records**

**400 000 authors
60 % of the catalogue
Periodics**

External links
(LCSH,
Dbpedia...)

Lab (Atelier)

Pages about places

Pages about dates

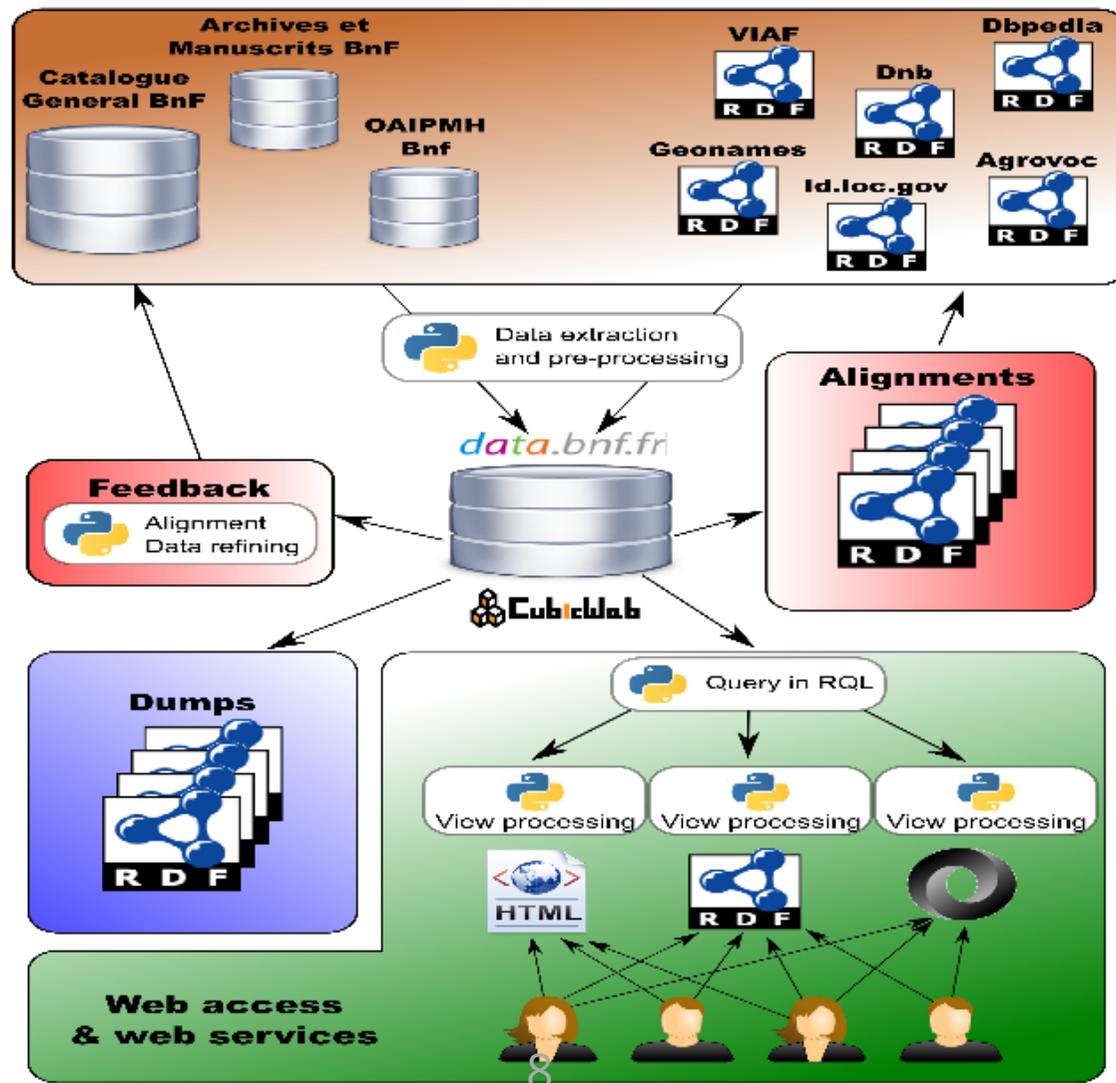
FRBR « Expressions »

Documents
about an
author/work

Web legal deposit

ISNI
reliures.bnf.fr, bp16.bnf.fr

**IFLA
2014
LYON**



Alignements

- Outil Nazca (calcul de similarités)
 - notice bibliographique <-> notice d'autorité titre
 - regroupe ou crée notices biblio et autorité
 - liens avec bases externes
- Réutilisation de données
 - règles métiers ou identifiant ARK
 - VIAF et autres alignements existants

identifiant ark

ark:/12148/cb11916599s

limit (0 if all)

500

minimum number of words by feature

1

maximum number of words by feature

4



Work ark:/12148/cb12186985d - Apologie de Raymond Sebon (4 alignments)

most relevant tokens: [apologie](#), [apologie raymond](#), [raymond](#)

Conserver	Confiance	Ark	Doctype	Catalog	Titre de manifestation
<input checked="" type="checkbox"/>	93%	ark:/12148/cb32461221b	livres	catalogue	Apologie de Raimond de Seb factice de Jamet et des notes [manifestation]
<input checked="" type="checkbox"/>	93%	ark:/12148/cb324612231	livres	catalogue	Montaigne. Apologie de Raimond de Sebon
<input checked="" type="checkbox"/>	93%	ark:/12148/cb33103838k	livres	catalogue	Montaigne. Apologie de Raimond de Sebon
<input checked="" type="checkbox"/>	93%	ark:/12148/cb37038207v	livres	catalogue	Apologia de Ramon Sibiudaj

generate json alignment rules generate csv alignment rules

Work ark:/12148/cb119426809 - Les essais (36 alignments)

most relevant tokens: [essais](#), [essais francais](#)

Conserver	Confiance	Ark	Doctype	Catalog	Titre de manifestation
<input checked="" type="checkbox"/>	97%	ark:/12148/cb30276252d	livres	catalogue	Supplément aux " Essais " de Montaigne
<input checked="" type="checkbox"/>	97%	ark:/12148/cb306511181	livres	catalogue	Principaux chapitres et extraits de Les Essais de Montaigne
<input checked="" type="checkbox"/>	97%	ark:/12148/cb30967153q	livres	catalogue	Livre des Essais de Montaigne, édition de 1843. [Voir tous les détails]
<input checked="" type="checkbox"/>	97%	ark:/12148/cb309673100	livres	catalogue	Reproduction en photocopie de Les Essais de Montaigne appartenant à la bibliothèque de Fortunat Strowski, ... L

Feuille de route data.bnf

- 100% du catalogue
- mise à jour toutes les 2 semaines
- réintégration des alignements **dans** le catalogue
- ajout de modèles plus simples (?)
- data.bnf.fr/atelier pour avant-premières

Enjeux data.bnf selon la BnF

- Visibilité et référencement
- Interopérabilité
- Fiabilité grâce à la traçabilité des URLs
- Réutilisation souple par des tiers



Cousteau, Nicolas (imprimeur-libraire, 14..?-15..)

1 | 2 | 3 | 4 | 5 | 6 | 7 Suivant »

[montrer 122 résultats](#)

Auteur	Titre	Date
Jacques de Voragine (1228?-1298)	La Legende doree ou Vie des saintz et saintes	1506
	[Bible. N.T.. français. 1524. Lefèvre d'Étaples] Les choses contenues en ce present liure... La S. Euangile selon S. Matthieu. La S. Euangile selon [...]	1524



Nicolas Cousteau (imprimeur-libraire, 14..?-15..)



Pays : **France**

Naissance : **14..**

Mort : **15..**

Activité commerciale : Imprimeur libraire

Note : Fils et successeur de Gilles Cousteau. - Au début de son exercice, travaille en association avec son frère Antoine sous la raison : "Antoine et Nicolas les Cousteaux"

Autre forme du nom : Nicolas Couteau (imprimeur-libraire, 14..?-15..)

ISNI : **ISNI 0000 0000 3361 7913**

Ses activités
(129 documents)

Documents à propos
de cet auteur

Pages dans data.bnf.fr
(2 pages)

Sources
et références

Recherche dans tout le site



Imprimer la page | Télécharger les données en rdf



plat supérieur

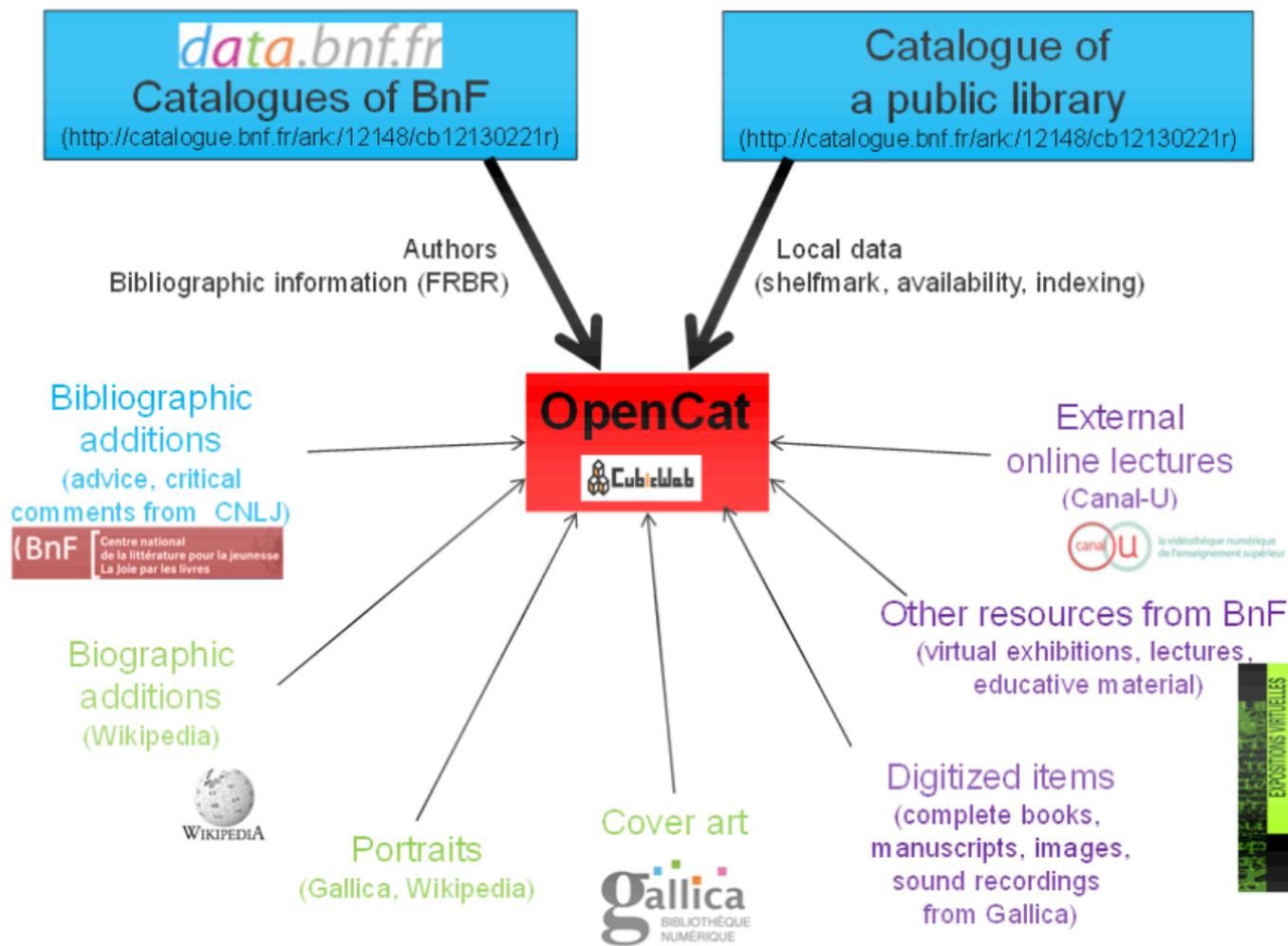
Reliure en veau roux à décor de plaques estampées à froid au nom d'André Boule, Paris, atelier d'André Boule, vers 1513-1523

Paris. Bibliothèque nationale de France. Réserve des livres rares. RES- E*- 68 (1) et (2)

Mots clés Reliure à décor • Reliure de libraire ♦ Plein cuir ♦ Veau ♦ Plaques figuratives ♦ Estampage à froid • Fers • Plaques • Roulettes ornées

Dimensions : 289 x 208 x 57 mm







Domenico SCARLATTI (1685-1757)

Naissance	1685, Naples (Italie)
Décès	1757, Madrid
Activité	compositeur

Ressources externes

bnf.fr

Pour plus d'information sur cette personne voir data.bnf.fr →

dbpedia

from http://en.wikipedia.org/wiki/Domenico_Scarlatti →

“(Giuseppe) Domenico Scarlatti est un compositeur baroque et claveciniste virtuose italien, né à Naples le 26 octobre 1685, mort à Madrid le 23 juillet 1757. Né à l'époque de Johann Sebastian Bach, Domenico Scarlatti passe la première partie de sa vie dans le sillage et à l'ombre de son père Alessandro Scarlatti, musicien très renommé. Claveciniste virtuose, compositeur d'opéras, musicien de cour ou d'église, il ne parvient pas, cependant, à se fixer durablement et à faire carrière dans une des villes de ses pérégrinations : Naples, Rome, Florence, Venise... Quelques années — [...]lire plus ▾”.

Enregistrements (1)

William CHRISTIE dirige Les Arts Florissants

Lamentazione CD_50999 0709072 1 - copie [écouter dans un player](#)

catégories sélectionnées :

paysage

Permalink

MUSÉE DES AUGUSTINS
MUSÉE DES BEAUX-ARTS DE TOULOUSE

Musée des Augustins

COURBET Gustave



Domaines d'activité
Écoles
Identique à

peinture
France
<http://data.bnf.fr/ark:/12148/cb120519068>

External resources

http://dbpedia.org/resource/Gustave_Courbet



Gustave Courbet

Jean Désiré Gustave Courbet (10 June 1819 – 31 December 1877) was a French pair French painting. The Realist movement bridged the Romantic movement (character: Eugène Delacroix) with the Barbizon School and the Impressionists. Courbet occupie painting as an innovator and as an artist willing to make bold social commentary in hi lived in freedom; let me end my life free; when I am dead let this be said of me: 'He be to no academy, least of all to any régime except the régime of liberty.' http://dbpedia.org/resource/Category:L%C3%A9gion_d'honneur_refusals, http://dbpedia.org/resource/Category:19th-century_French_people, http://dbpedia.org/resource/Category:1819_births, http://dbpedia.org/resource/Category:1877_deaths, http://dbpedia.org/resource/Category:1877_deaths, http://dbpedia.org/resource/Category:1877_deaths



Linked Data et bibliothèques

- <http://data.bnf.fr> et bibliothèques
- <http://data.europeana.eu>
- <http://datos.bne.es>, <http://dnb.de>, <http://libris.kb.se>
- Sudoc IdRef
- VIAF (Virtual International Authority File)
- etc

Provenance des jeux de données

- généré par ?
- utilisé par ?
- date de début / fin ?
- dérivé de ?
- licence ?

Maturité des données ouvertes

1. Données disponibles sur le Web sous une licence ouverte
2. + format structuré (CSV, XML, Excel, mais pas PDF ou image)
3. + format non-propriétaire (CSV plutôt que Excel)
4. + respect des standards RDF (identification par des URLs, vocabulaire défini: pas CSV, etc.)
5. + liens vers d'autres données pour fournir du contexte (ontologies) et enrichir (autres jeux de données)

interopérabilité

le Web

Qu'est-ce que le web ?



Réseau mondial de données construit sur la base d'un ensemble de **standards d'interopérabilité** (protocoles, formats, vocabulaires, modèles, interrogation), qui s'appuie sur les **URLs** pour identifier de manière unique les données élémentaires et les **lier** entre elles.

Le Web passe à l'échelle

Liens permettent collaboration à grande échelle

Création d'identifiants est **décentralisée**

(Domain Name System est délégué)

Incohérence globale ("monde ouvert") permet **parallélisation**

Attention à ne pas confondre

- **web sémantique :**

données utilisables par un programme PAS trait. auto. des langues

- **RDF et vocabulaires :**

standardiser les échanges de données PAS les outils et le stockage

- **Un et un seul Web :**

un protocole (HTTP) et plusieurs standards pour l'échange de données

(HTML, RDF, OWL, JSON, DC, FOAF, SKOS, etc.)

Recentrage sur les données

- rassembler / collecter
- aligner / normaliser
- enrichir / annoter / classer
- définir un pivot métier avec des données de référence
- lier les données entre elles

Urbanisation des données

libre	privé
externe	interne
générique	spécifique
statique	dynamique

Avenir probable

- biens communs
- spécialisation territoire / langue / thème / domaine
- interopérabilité / décentralisation
- multiplicité des réutilisations

*“ Les applications
passent,
les données
restent.”*