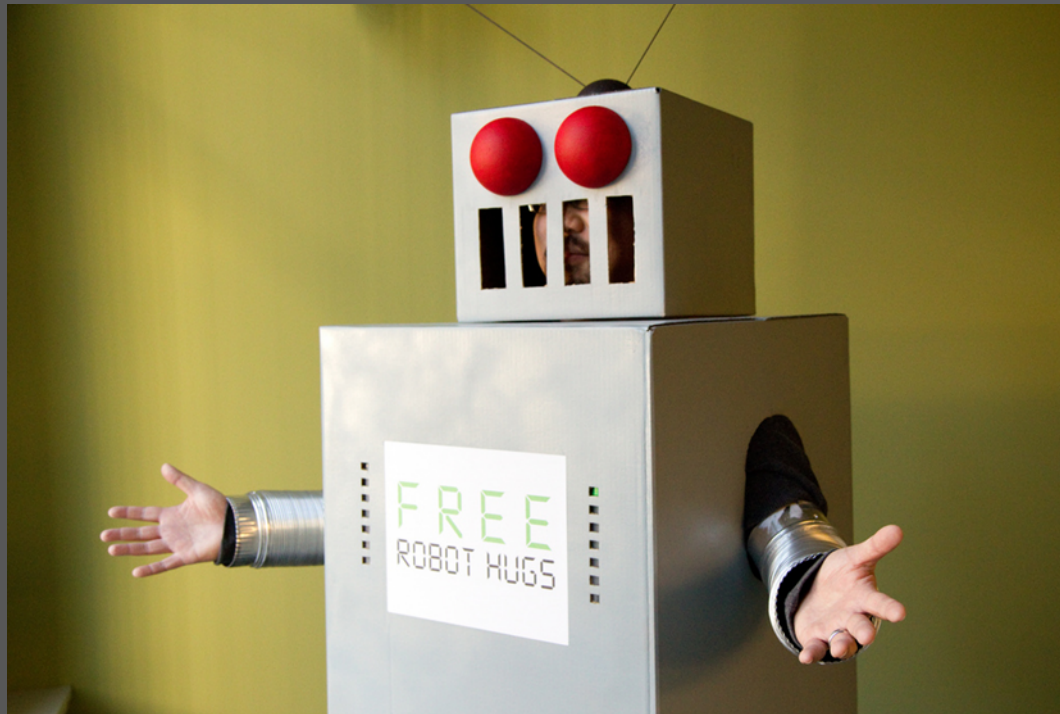


LE LINKED OPEN DATA, UNE REPONSE AUX NOUVELLES PROBLEMATIQUES DES CATALOGUES ?



Documentation mathématique, RBNBM, Marseille, 17 novembre 2015

{BnF

Gildas ILLIEN, Département de l'Information
bibliographique et numérique

Introduction



data.bnf.fr

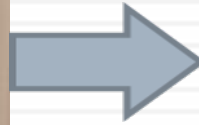
Plan



- Motivations et principes du *Library Linked Data* vu depuis la BnF
- Data.bnf.fr et illustrations dans le domaine des mathématiques
- Perspectives et questionnements

Motivations et principes clés

Le Library Linked Data vu de la BnF



Les catalogues de bibliothèque et le Web : en bref

- **Episode 1** : le catalogue contre le Web
- **Episode 2** : le catalogue **sur** le Web
- **Episode 3** : le ~~catalogue~~ **dans** le Web
les « données » ?



Notre place est ici



Linked Open Data (LOD), c'est-à-dire?

□ **LINK = Lier les données :**

- pour permettre leur interopérabilité et les placer au coeur des usages numériques
- en transformant et en exposant des données structurées par et pour les bibliothèques (MARC) selon les standards du Web sémantique (identifiants, référentiels, RDF)

□ **OPEN = Ouvrir les données:**

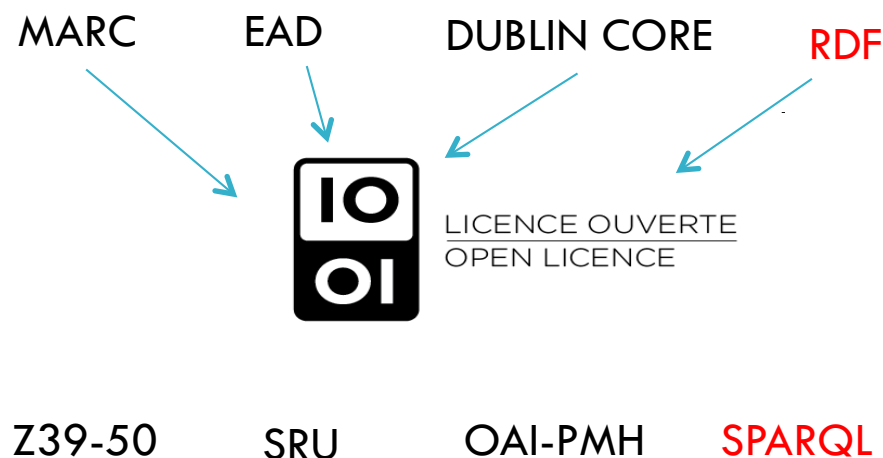
- Techniquement
- Juridiquement
- + un nouvel état d'esprit?



LICENCE OUVERTE
OPEN LICENCE

L'ouverture des métadonnées BnF

- **Périmètre** : Toutes les **métadonnées** dont la BnF est productrice, ou issues d'un régime compatible (OCLC...)
- **Licence** : licence ouverte de l'Etat (réutilisation libre et gratuite, avec obligation d'attribution), conçue par Etalab.



Pour quoi faire?

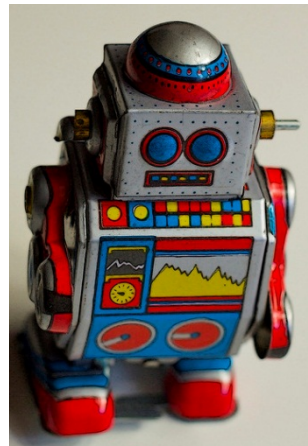
- ❑ **EXPOSER - Être plus visible**
 - L'enjeu du référencement par les moteurs de recherche
- ❑ **REGROUPER - Être plus cohérent**
 - Faciliter la fédération de silos de données de nature, de provenance et de structure différentes
- ❑ **PARTAGER - Être économe et généreux à la fois**
 - Prendre acte de la nouvelle économie des données
 - Se lier à d'autres données de confiance
 - Encourager les réutilisations – par les autres bibliothèques, mais pas seulement.
- ❑ **INNOVER – Être à l'avant-garde du web sémantique**
 - Garantir la maîtrise publique de référentiels nationaux et revisiter la mission de la Bibliographie nationale française à travers le web sémantique.

data.bnf.fr: une vision, trois objectifs, deux cibles



par des humains

Être visible



par des machines

Être lisible



Être réutilisable

Attention à la pensée magique : le LOD s'inscrit dans une transition plus globale

- ❑ Le LOD c'est d'abord un principe de large diffusion et d'encouragement à la réutilisation...
- ❑ Qui rencontre un important chantier de transformation de la structure des données (FRBR)...
- ❑ Lequel s'articule avec l'évolution des pratiques de catalogage (RDA)...
- ❑ Mais n'est pas la panacée : les bibliothèques ont d'autres chats à fouetter par ailleurs, et d'autres types de données à traiter.



Petit détour pédagogique



Le format = le véhicule informatique

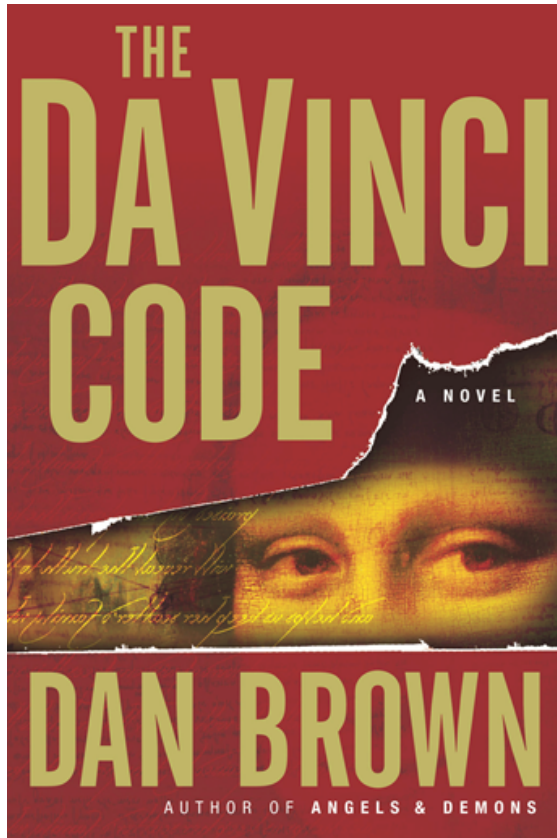


MARC
(BIBFRAME)

Production/diffusion

Riches/pauvres

Le code = les règles de description



Normes AFNOR

AACR2

RDA

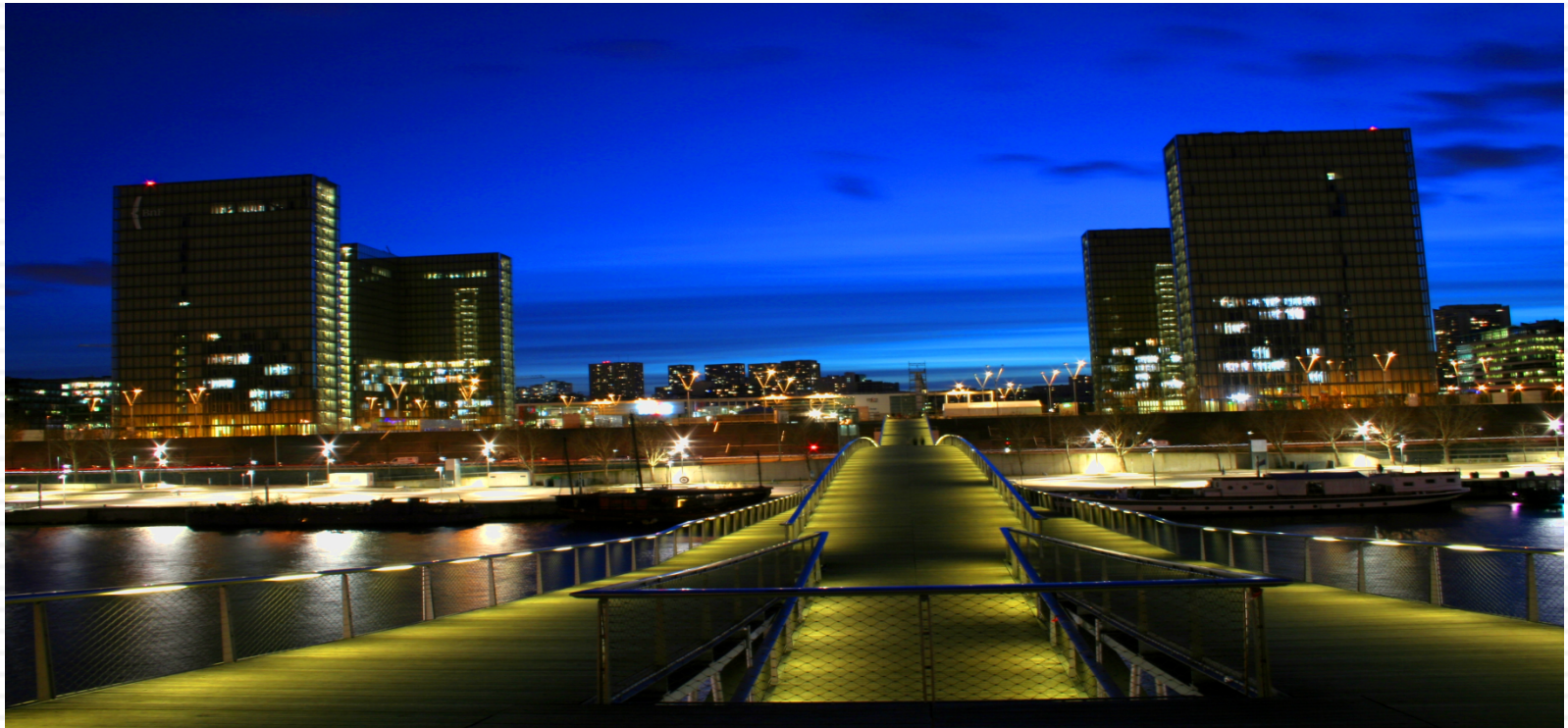
Le modèle = l'architecture de l'information



Functional Requirements for Bibliographic Records

- Entités/autorités
- Regroupements
- Mieux intégrer le numérique
- Parler la langue des internautes

Data.bnf.fr

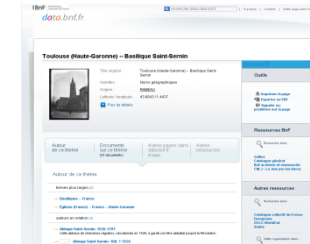


data.bnf.fr : Principes généraux

Collections numérisées (+3M)



pages Web pour les humains



BnF Archives & Manuscrits



données structurées pour les machines

```
<rdf:RDF>
- <rdf:Description rdf:about="http://data.bnf.fr/ark:/12148/cb119344774">
  <skos:editorialNote xml:lang="fr">GLE</skos:editorialNote>
  <skos:broadener rdf:resource="http://data.bnf.fr/ark:/12148/cb12194515v7">
  <skos:broadener rdf:resource="http://data.bnf.fr/ark:/12148/cb120093384">
  <owl:sameAs rdf:resource="http://stich.es.vu.nl/vocabularies/frameau/ark:/12148/cb119344774">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept">
  <rdfs:seeAlso rdf:resource="http://catalogue.bnf.fr/ark:/12148/cb119344774">
  <skos:inScheme xml:lang="fr">Noms géographiques</skos:inScheme>
  <skos:prefLabel xml:lang="fr">Toulouse (Haute-Garonne) -- Basilique Saint-Sernin</skos:prefLabel>
</rdf:Description>
</rdf:RDF>
```

Catalogue général (14 M)

{BnF
catalogue
général



Télécharger les données

- Télécharger en RDF
- Télécharger en Json

La récupération de données dans data.bnf.fr

A LA SAMARITAINE

RUE
DU
PONT-NEUF
13&5

INAUGURATION
DES
AGRANDISSEMENTS
EXPOSITION



LES COMPTOIRS
de Costumes et de Confections
DE LA SAMARITAINE
sont les plus BEAUX,
les plus VASTES,
les mieux ASSORTIS
et ceux qui vendent
le MEILLEUR MARCHÉ
DE PARIS

23^f COSTUME *pure Laine*
TISSU BEIGE
Jupe haut plissé, Echarpe laveuse

18^f VISITE *beau Mérinos double*
GARNIE EFFILÉ,
tout Canaque et Passementerie

LITH. LE DÉTACHÉ, R. ... TEMPLE, N. PARIS

<http://gallica.bnf.fr/ark:/12148/btv1b9005004h>

W 008495

Oa

Rouleau



Que trouve-t-on
dans data.bnf.fr?

Que trouve-t-on dans data.bnf.fr ?

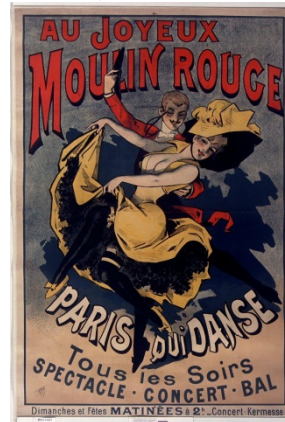
Typologie

des personnes



des œuvres

lues, traduites, interprétées
éditées, manuscrites, numérisées

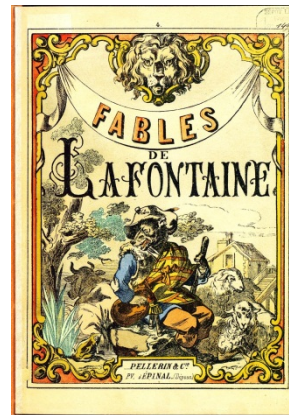


des dates



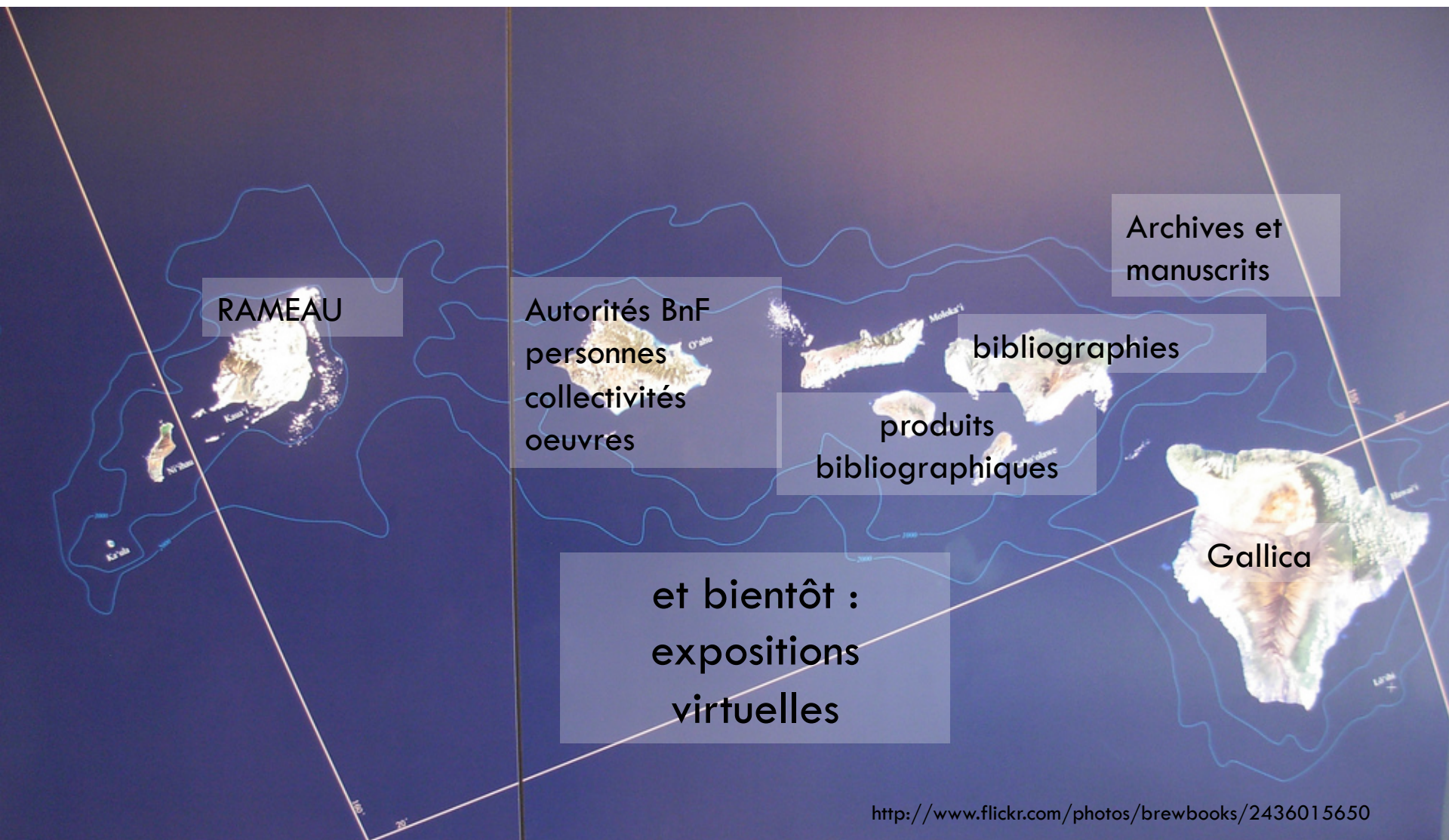
des lieux

des organisations



des sujets

Quelles ressources BnF ?



DEMO POUR MATHEUX

- **Depuis Google:**
- Syntaxis mathematica claudes ptolémée
- diophante les arithmétiques
- pierre de fermat auteur
- Applications de géométrie descriptive adhémar

- **Quelques belles pages :**
- des liens à BAM, Gallica, expositions virtuelles
- <http://data.bnf.fr/12013155/averroes/>
- http://data.bnf.fr/11920750/claude_ptolemee/
- des liens à Gallica
- http://data.bnf.fr/11920108/henri_poincare/
- plus moderne :
- http://data.bnf.fr/12279721/andre_weil/ pas de documents numérisés, mais lien aux bibliographies de la BnF

- **exemples d'oeuvres numérisées :**
- <http://gallica.bnf.fr/ark:/12148/btv1b8625611h> théorie des probabilité de laplace
- <http://gallica.bnf.fr/ark:/12148/bpt6k62145354> oeuvre complète de Pierre de Fermat

Récupérer les données en RDF : comment ?

LA GRANDE GUERRE

N° 24



MODÈLE DÉPOSÉ

LES VENDANGES DE 1914

IMP., 13, QUAI D'AJACCIO, PARIS

NOS VIGNERONS CHAMPENOIS ONT VENDANGÉ SOUS UNE PLUIE DE FER

Nos vignerons ont fait preuve d'un rare courage : ils ont vendangé malgré un bombardement intense et continu.

Lors de l'invasion de notre pays, les Allemands épargnèrent nos vignes avec la certitude que le vin serait pour eux ; or nos troupes ont chassé si rapidement l'ennemi, qu'il n'a pas eut le temps de

saccager les récoltes et se retirant. Il continua néanmoins à bombarder sans relâche toute la région rémoise, mais les vendanges s'effectuèrent avec entrain et dans bien des endroits l'on put voir des vendangeuses distribuer quelques belles grappes de raisins à nos troupes. Le champagne de 1914 sera un vin exceptionnellement bon et les Allemands n'en auront pas. (Novembre 1914).

<http://gallica.bnf.fr/ark:/12148/btv1b10213439b>

Les trois modes de récupération du RDF

I. Tout en bloc



<https://www.flickr.com/photos/caliterra/3145971395>

télécharger tout
data.bnf.fr

[http://data.bnf.fr/
semanticweb#Ancre2](http://data.bnf.fr/semanticweb#Ancre2)

- en totalité
- par type de ressource

Les trois modes de récupération du RDF

II. Pour chaque page BnF



<https://www.flickr.com/photos/jfgornet/12115178303>

Exemples :

[Paul Claudel](#)

Tête d'Or

- [l'œuvre originale](#)
- le [spectacle de Jean-Louis Barrault](#)

[Le Vésinet](#) (lieu)

Les trois modes de récupération du RDF

III. Comme vous voulez



Exemples :

- œuvres en breton : [requête](#)
- images numérisées sur des lieux du Morbihan : [requête](#)
- les œuvres les plus numérisées : [requête](#)
- les 100 auteurs les plus étudiés : [requête](#)

[Guide SPARQL](#) data.bnf.fr -
exemples de requêtes

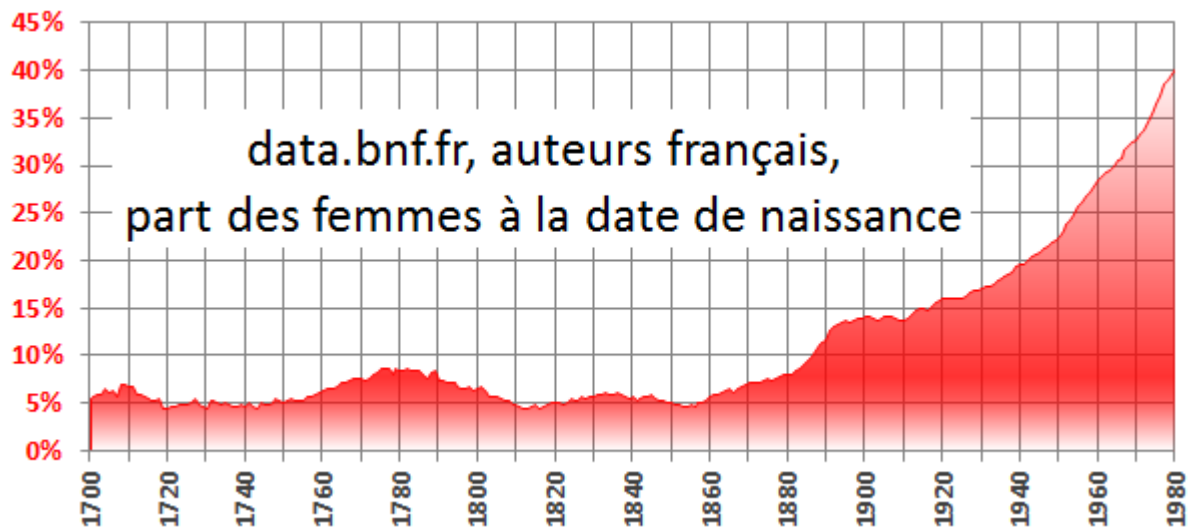
Résultats et actualité de data.bnf.fr

□ Chiffres clés

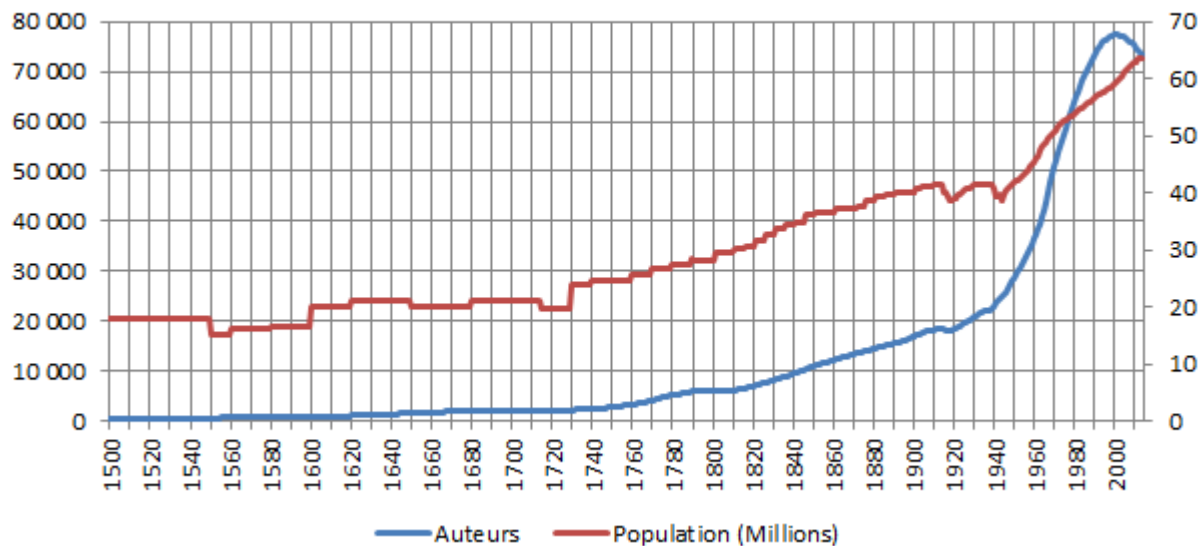
- Une fréquentation en hausse constante (13 000 v/j)
- 80% viennent des moteurs de recherche.
- 80% des catalogues exposés. Env.100% mi 2016.

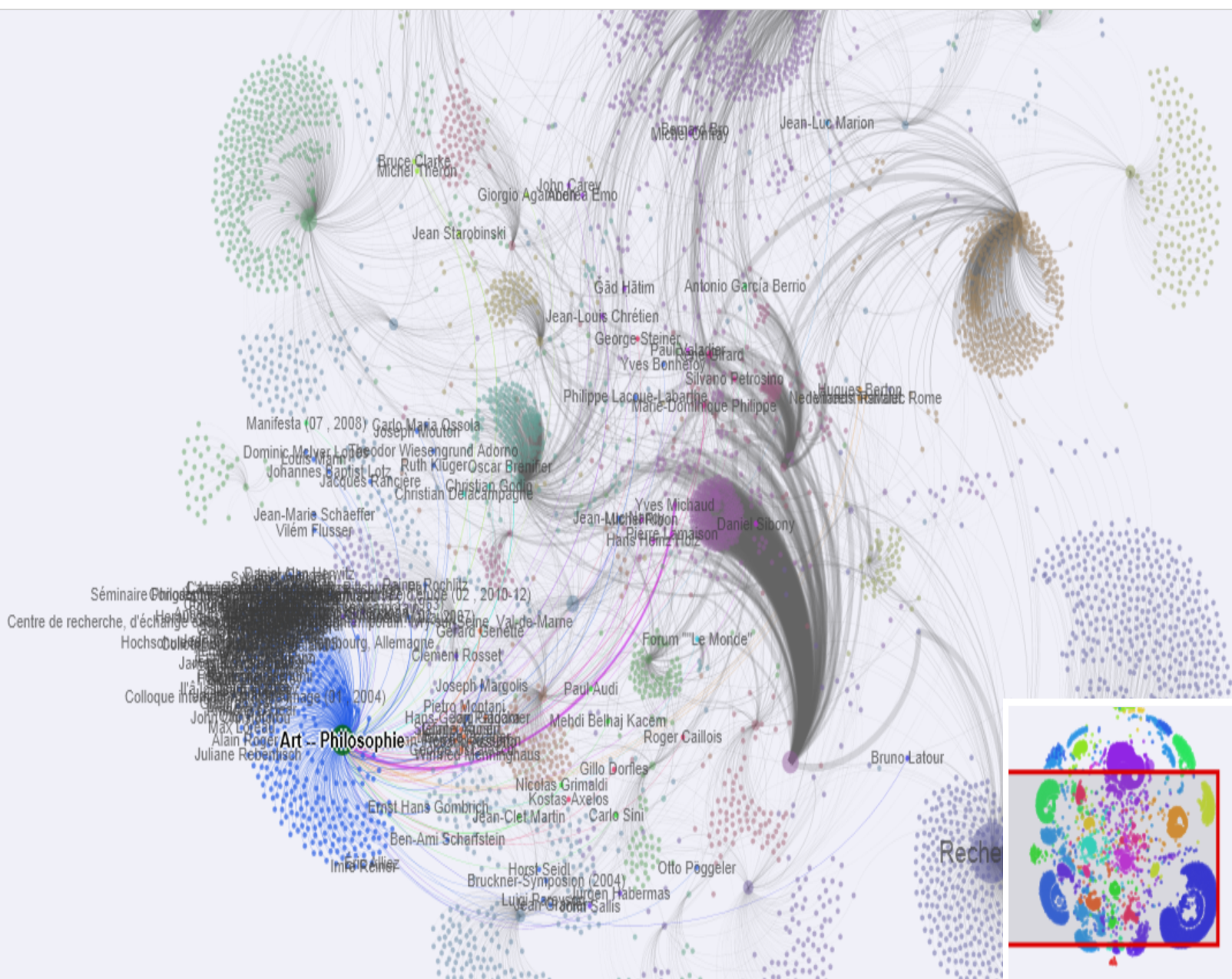
□ Axes de développement en discussion

- ▣ FRBRisation (machine à retraiter les données source)
- ▣ Datavisualisation et fouille de données (metadata mining)
- ▣ Accompagner les réutilisations
- ▣ Approfondir la coopération avec d'autres tiers de confiance



France 1500–2015, auteurs et population totale

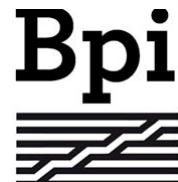




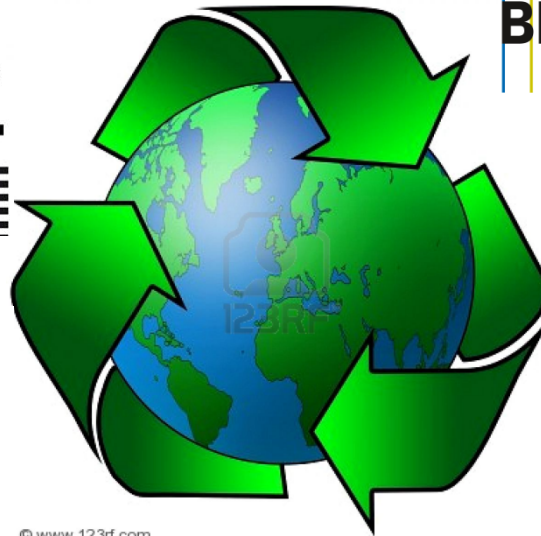
Navigation sidebar containing icons for: Home, Search, Zoom in (+), Zoom out (-), and a magnifying glass icon.

QUELQUES EXEMPLES DE REUTILISATION

abes
agence bibliographique de l'enseignement supérieur



fevis



© www.123rf.com



WIKIPEDIA

Perspectives et questionnements



Grands enjeux et Points de passage

- La FRBRisation, pour de vrai?
- RDA.FR et le programme Transition bibliographique
- L'Open (Meta)Data : l'ouverture juridique et technique et son impact sur le marché de la métadonnée
- Maîtriser et expérimenter les technologies et standards du web sémantique : c'est vite dit...
- Focus sur les données d'autorité. Ex. de VIAF
- Focus sur les identifiants. Ex. de ISNI
- Décloisonnements nationaux et internationaux.

Le catalogue est mort, vive le catalogue

- ❑ Les métadonnées de qualité resteront la clé de voûte du fonctionnement et de la présence des bibliothèques, leur marque de confiance, même si elles deviennent transparentes, cachées (cf. Google).
- ❑ On en aura toujours besoin pour fonctionner. Et pour être vus, utilisés.
- ❑ C'est vrai qu'il faut arrêter de dupliquer les efforts.
- ❑ C'est vrai qu'on peut et qu'on doit dériver ou se lier.
- ❑ Mais il faut quand même que quelqu'un quelque part continue de cataloguer correctement.
- ❑ Ce capital (et cette compétence) ne doivent pas échapper complètement aux acteurs publics. Enjeu identitaire majeur.
- ❑ La maîtrise de métadonnées ouvertes et interopérables est au cœur de toutes les stratégies d'innovation et de transition.
- ❑ La question des systèmes se pose dans une autre temporalité. Ca change tout le temps. La porosité entre systèmes et données est un enjeu majeur.

Nouvelles frontières, nouvelles coopérations?

LES GENS



LES BIBLIOTHEQUES



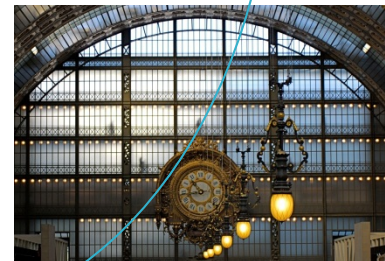
LES EDITEURS



LA RECHERCHE



LES MUSEES



LES ARCHIVES

En savoir plus...



SUR LE SITE DE LA BNF :

[http://www.bnf.fr/fr/professionnels/
innov_num_web_donnees.html](http://www.bnf.fr/fr/professionnels/innov_num_web_donnees.html)

SUR LE SITE DATA.BNF.FR :

<http://data.bnf.fr>



MERCI DE VOTRE
ATTENTION !

{ BnF

gildas.illien [AT] bnf.fr