

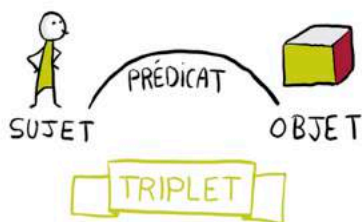
# Un triplestore pour Persée



## Un triplestore, qu'est-ce que c'est ?

Un endroit (virtuel) où on a rangé des millions de toutes petites informations : des triplets.

Un triplet c'est une information descriptive en trois parties : un sujet, un prédicat (une propriété), un objet.



Par exemple "La Distinction" "a pour auteur" "Pierre Bourdieu".

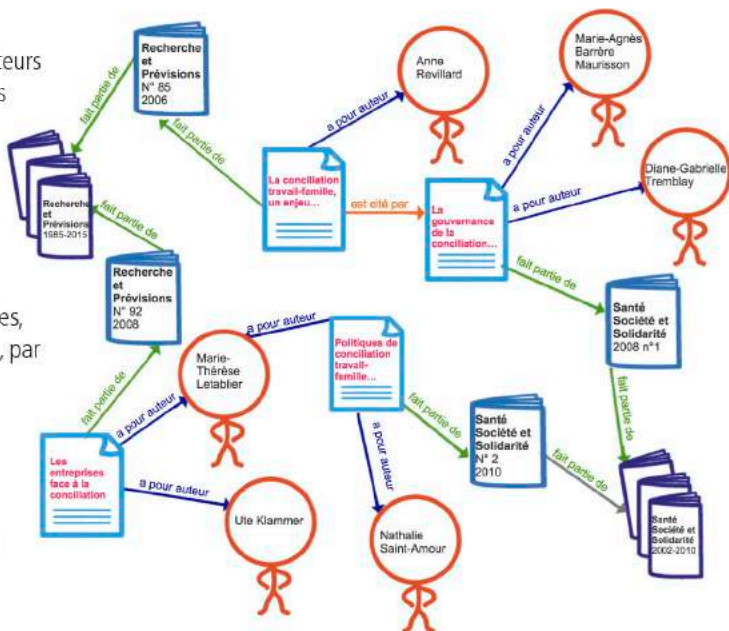
Chaque personne est définie par une dizaine de propriétés (nom, prénom, dates de naissance et de décès, etc.), chaque document par une vingtaine, les illustrations, fascicules, collections, monographies et concepts sont également décrits. A ces triplets s'ajoutent les relations entre entités, les 700 000 documents et 115 000 personnes du

fonds de Persée donnent donc lieu à 30 millions de triplets !

On assigne des identifiants uniques à tous les objets décrits (les documents, les personnes...) et à tous les prédicats. De cette façon, toutes les informations concernant un "objet" peuvent être regroupées, et à chaque fois qu'on utilise un prédicat on parle strictement de la même chose.

Par exemple, ici, nous avons quatre articles liés à leurs auteurs et liés entre eux par différents prédicats : "est cité par", "a pour auteur", "fait partie d'un fascicule" qui "fait partie d'une collection". Ce graphe est volontairement simple. On aurait également pu lier les entités par des dates, par des éditeurs scientifiques, par des lieux de publication...

Ces entités et leurs liens forment un graphe. Sur data.persee.fr, c'est ce graphe de 30 millions d'informations liées que l'on interroge.



## Qu'est-ce qu'on y cherche ?

On y cherche des informations sur les documents, les personnes, les illustrations, les fascicules, les revues, les concepts qui font partie du fonds de Persée.

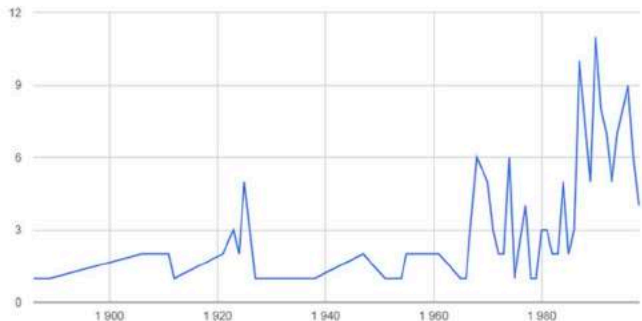
On ne cherche pas le texte intégral, celui-ci est sur Persee.fr. Mais on peut y accéder en recherchant les identifiants des articles et en cliquant sur ceux-ci. Les recherches sont personnalisées : on cherche ce qui nous intéresse et rien que ce qui nous intéresse.

## Exemples d'utilisation :

On peut chercher des listes d'entités partageant un certain nombre de propriétés : tous les articles parus entre 1850 et 1950 dans des revues d'anthropologie dont le titre contient "mariage", on peut ainsi se **créer un corpus de recherche**.

Identifiant	Date de Publication	Titre
<a href="http://data.persee.fr/doc/corr_0000-0002_1863_cor_2_1_922_12_0068_000">http://data.persee.fr/doc/corr_0000-0002_1863_cor_2_1_922_12_0068_000</a>	1863	Proposition de Jean-Baptiste Colbert (contrôleur général des
<a href="http://data.persee.fr/doc/crai_0065-0536_1865_num_9_1_67046#Web">http://data.persee.fr/doc/crai_0065-0536_1865_num_9_1_67046#Web</a>	1865	Les mariages, 1571-72
<a href="http://data.persee.fr/doc/bmsap_0301-8644_1871_num_6_1_4472#Web">http://data.persee.fr/doc/bmsap_0301-8644_1871_num_6_1_4472#Web</a>	1871	Parenté des époux ou mariages consanguins
<a href="http://data.persee.fr/doc/bmsap_0301-8644_1872_num_7_1_4508#Web">http://data.persee.fr/doc/bmsap_0301-8644_1872_num_7_1_4508#Web</a>	1872	Sur le mariage au point de vue anthropologique
<a href="http://data.persee.fr/doc/bmsap_0301-8644_1877_num_12_1_3225#Web">http://data.persee.fr/doc/bmsap_0301-8644_1877_num_12_1_3225#Web</a>	1877	Rapport sur le mémoire de M. Lacassagne sur les mariages
<a href="http://data.persee.fr/doc/bmsap_0301-8644_1880_num_3_1_3307#Web">http://data.persee.fr/doc/bmsap_0301-8644_1880_num_3_1_3307#Web</a>	1880	De l'influence du mariage sur la tendance au suicide
<a href="http://data.persee.fr/doc/roma_0035-8029_1880_num_9_36_6549#Web">http://data.persee.fr/doc/roma_0035-8029_1880_num_9_36_6549#Web</a>	1880	Un mariage dans le Haut-Foréz. Usages et chants
<a href="http://data.persee.fr/doc/bmsap_0301-8644_1891_num_2_1_7560#Web">http://data.persee.fr/doc/bmsap_0301-8644_1891_num_2_1_7560#Web</a>	1891	Le mariage aux Indes
<a href="http://data.persee.fr/doc/anami_0003-4398_1895_num_7_28_6670#Web">http://data.persee.fr/doc/anami_0003-4398_1895_num_7_28_6670#Web</a>	1895	VII. Extrait d'un contrat de mariage de 1512, à propos des noi
<a href="http://data.persee.fr/doc/phloui_0776-5541_1895_num_2_6_1418_t1_0226">http://data.persee.fr/doc/phloui_0776-5541_1895_num_2_6_1418_t1_0226</a>	1895	Edouard Westeemark. Origine du mariage dans l'espèce hu

On peut chercher des volumes : le nombre d'articles par an, parus entre 1880 et 1998, dont le titre contient le mot "nation". On peut ainsi **visualiser des tendances**.



On peut chercher des liens entre des personnes ou des documents : les documents qui citent ou sont cités par les articles de telle personne. On peut ainsi **mesurer l'influence** d'un auteur ou d'une revue, l'impact d'un article sur sa discipline.

On peut aussi **télécharger un jeu de données** : une partie thématique ou l'intégralité du graphe, que l'on pourra étudier, manipuler, réutiliser sur son propre ordinateur.

 <p><b>Bibliothèque de l'École des chartes</b></p> <p>Format : <b>RDF</b> Date de génération : 10/01/2017 Période : 1840-2007 Volumétrie : 212 numéros - 15758 documents - 1703 personnes - 3,69 Mo</p>	 <p><b>Bibliothèque des Écoles françaises d'Athènes et de Rome</b></p> <p>Format : <b>RDF</b> Date de génération : 10/01/2017 Période : 1976-1998 Volumétrie : 26 numéros - 26 documents - 48 personnes - 210,98 Ko</p>	 <p><b>Bulletin de l'École française d'Extrême-Orient</b></p> <p>Format : <b>RDF</b> Date de génération : 10/01/2017 Période : 1901-2000 Volumétrie : 109 numéros - 3552 documents - 694 personnes - 1,24 Mo</p>	 <p><b>Bulletins et Mémoires de la Société d'Anthropologie de Paris</b></p> <p>Format : <b>RDF</b> Date de génération : 10/01/2017 Période : 1864-2009 Volumétrie : 304 numéros - 7029 documents - 1864 personnes - 2,46 Mo</p>
--	--	---	--

## Comment on s'en sert ?

En général, pour chercher dans un triplestore, il faut connaître un langage informatique de requête : le Sparql.



## Sparklis

Pour interroger le triplestore de Persée, nous vous proposons une solution plus accessible : Sparklis, qui permet de faire des requêtes en langage naturel. La requête en Sparql est générée par le programme, vous pouvez l'afficher si vous le souhaitez.

give me every document  
that is the part of an issue that is the part of a journal that has a title  
and whose aut has as a name Pierre Bourdieu  
and that has a date of print publication >  
and for each title  
give me the highest-to-lowest number of document  
and for each title  
give me the list of date of print publication

**Sparklis suggestions to refine your query**

The current focus is **the document** (click on different parts of the query to change it)

matches all of 1 | matches all of 2 | matches all of 3 | **OK**

every document > (116)  
every manifestation (116)  
that is the focus of ... (116)  
every article (22)  
that has an access rights (31)  
that has a date of publication (51)  
that has an identifier (21)

Bensa Alban, Bourdieu Pierre. Quand les Canoques prennent la parole. In: Actes de la recherche en sciences sociales. Vol. 56, mars 1985. L'anthropologie, pp. 69-85. [Online] © Bourdieu Pierre. Avenir de classe et causalité du probable. In: Revue française de sociologie, 1974, 15-1, nn. 3-42. [Online]

that is ...  
and ...  
or ...  
the highest-to-lowest ---  
the lowest-to-highest ---  
any ---  
a sample of ---  
the number of ---

37 concepts | 116 entities | 19 modifiers

**Results of your query**

results 1 - 10 of 116 | Show 10 | results

La requête en langage naturel se construit ici

On choisit les informations recherchées dans les menus déroulants

document	issue	journal	aut	title	date of print publication	number of document	list of date of print publication
1	Bourdieu Pierre, Driault Yvette. Le couturier et sa griffe contribution à une théorie de la magie. In: Actes de la recherche en sciences sociales. Vol. 1, n°1, janvier 1975. Hiérarchie sociale des	Actes de la recherche en sciences	Pierre Bourdieu	Actes de la recherche en	1975	95	1975 / 1976 / 1977 / 1978 / 1979 / 1980 / 1981 / 1982 / 1983 / 1984 / 1985 / 1986 / 1987 / 1988 / 1990 / 1991 / 1992 / 1993 / 1994 / 1995 /

Les résultats apparaissent sous forme de table téléchargeable

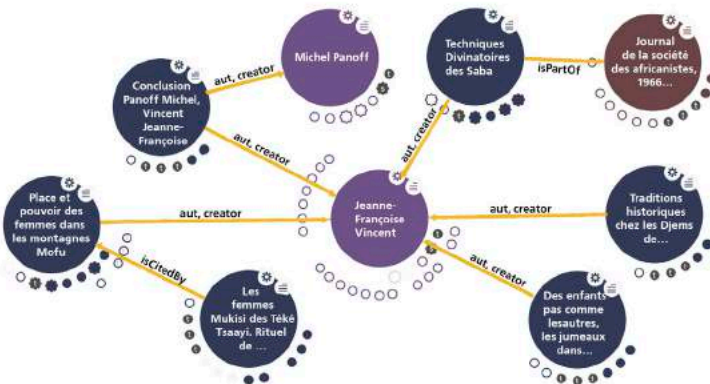
## Sparql endpoint

Pour ceux qui connaissent Sparql, une porte d'entrée directe est bien sûr également disponible.

Default Data Set Name (Graph IRI)  
http://data.persee.fr

Query Text  
`select distinct ?p where {?p a foaf:Person} LIMIT 100`

[About](#) | [Namespace Prefixes](#) | [Inference Rules](#)



## Visualisation

Vous pouvez aussi explorer le graphe avec un outil de visualisation. En partant d'une ressource - auteur ou document -, vous pourrez déplier les ressources liées et afficher la nature de ce lien.

## Data.persee.fr, c'est aussi...

Le triplestore de Persée s'intègre dans un site complet.

Vous y trouverez des contenus pédagogiques qui expliqueront plus en détail certaines notions techniques, des tutoriels pour vous apprendre à vous servir de Sparklis, un espace d'échanges avec la communauté scientifique et de l'aide pour formuler vos requêtes.

**DATA PERSÉE**

Persée en métadonnées

Sparklis Sparql endpoint Qu'est-ce qu'un triplestore ?

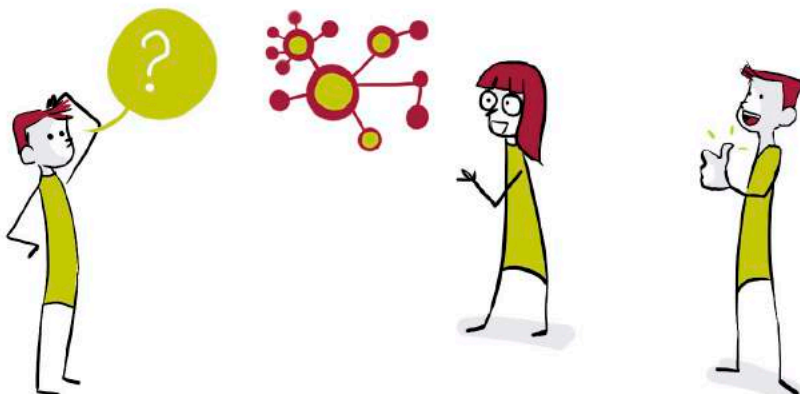
**Tutoriels**  
Apprenez, pas à pas, à rechercher des informations dans le triplestore de Persée.  
C'est parti !

**Explorer**  
Pour explorer nos métadonnées, optez pour Sparklis ou pour le sparql endpoint classique, téléchargez un jeu de données ou consultez les schémas de données.  
Consulter

**Comprendre**  
Apprivoisez Sparklis, l'interface qui vous fera écrire du Sparql sans le savoir, découvrez des outils de visualisation et soyez incollable sur les métadonnées et les triplestores.  
Consulter

**Partager**  
Partagez vos découvertes, demandez de l'aide, expliquez vos travaux et découvrez ce que font les autres utilisateurs de data.persee.fr de nos métadonnées.  
Consulter

Vous voulez en savoir plus ?  
Faire passer votre équipe au web de données ? Nous dispensons des **formations** à l'utilisation de notre triplestore, contactez-nous !



formation@persee.fr www.persee.fr data.persee.fr

