

Enseignement et formation en IA en ligne  
Retour d'expérience sur le MOOC  
Web de données et Web sémantique

Catherine Faron Zucker



[Accueil](#) • [Tous les cours](#) • [Web sémantique et Web de données](#)

# Web sémantique et Web de données

[Voir la page de présentation du cours dans Studio](#)

Les thématiques

Informatique

Numérique, technologie

Sciences pour l'ingénieur

## A PROPOS DU COURS

Ce cours vous propose de vous former aux standards du Web de données et du Web sémantique. Il vous présentera les langages qui permettent :

- de représenter et de publier des données liées sur le Web (RDF) ;
- d'interroger et de sélectionner très précisément ces données à distance et au travers du Web (SPARQL) ;
- de représenter des vocabulaires et de raisonner et déduire de nouvelles données pour enrichir les descriptions publiées (RDFS, OWL, SKOS) ;
- et enfin, de tracer et de suivre l'histoire des données (VOiD, DCAT, PROV-O, etc.).

## À QUI S'ADRESSE CE COURS ?

Ce cours s'adresse à des étudiants ou ingénieurs en informatique, notamment dans le domaine des systèmes d'information. Il peut être suivi en complément d'une formation classique aux technologies de base du Web.

[ Fabien Gandon, Olivier Corby, Catherine Faron Zucker, 2015 ]  
IDEFI uTOP Project - PIA ANR-11



# Conception

- plateforme Open edX FUN
- 7 semaines/sessions ~ 45 minutes
- 48 vidéos 5-7 minutes de cours (6-7 videos/semaine)
- 28 vidéos de démos (BBC, BNF, DBpedia, Corese, etc.)
- 7 sections de références avec des pointeurs
- 46 quizz entre les vidéos
- 7 long quizz & exercices de semaine (~1H30 tvl perso)
- 68 exercices indépendants au total (tous types de edX)
- 3 enseignants (4HM) et 2 ingénieurs (3HM)
- forum et animation : 60 heures

Vers un web de données liées

## Présentation de la semaine

### 1. Brève histoire du Web

Quiz



### 2. Principes architecturaux

Quiz



### 3. Standards et déploiement

Quiz



### 4. Séparation du fond et de la forme

Quiz



### 5. De la page à la ressource

Quiz



### 6. Principes des données liées

Quiz



### 7. La pile de standardisation

Quiz



## Compléments de cours

### Démos d'un Web de données liées

### Exercices de la semaine: un Web de données liées

Exercices de fin de semaine



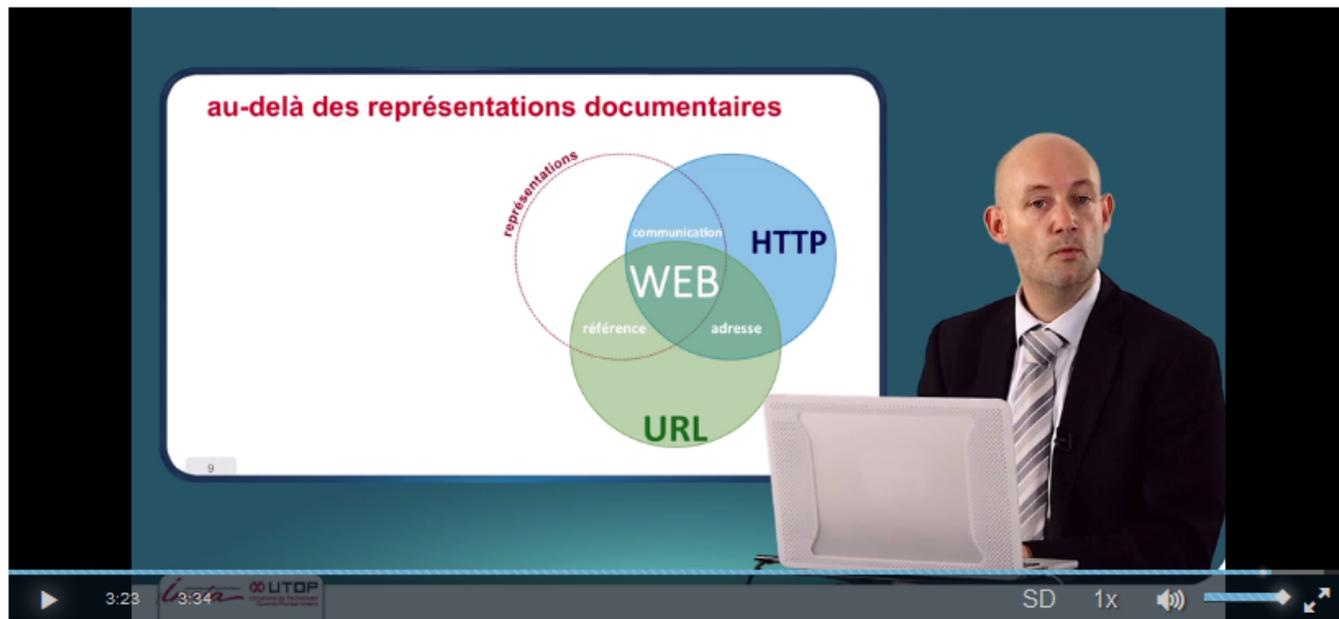
› Le modèle de données RDF

› Le langage de requête SPARQL

› Ontologies et schémas RDFS



VOIR L'UNITÉ DANS STUDIO



Télécharger la vidéo : Haute définition (720p) / Standard (512p) / Smartphone (320p)



vidéos de cours (~7min)

Vers un web de données liées

### Présentation de la semaine

#### 1.Brève histoire du Web

Quiz



#### 2.Principes architecturaux

Quiz



#### 3.Standards et déploiement

Quiz



#### 4.Séparation du fond et de la forme

Quiz



#### 5.De la page à la ressource

Quiz



#### 6.Principes des données liées

Quiz



#### 7.La pile de standardisation

Quiz



### Compléments de cours

#### Démos d'un Web de données liées

#### Exercices de la semaine: un Web de données liées

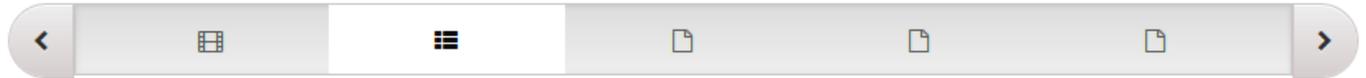
Exercices de fin de semaine



Le modèle de données RDF

Le langage de requête SPARQL

Ontologies et schémas RDFS



VOIR L'UNITÉ DANS STUDIO

QUIZ (1 point possible)



### LA PILE DE STANDARDISATION

Dans l'architecture Web sémantique, quels changements fait-on pour décrire des ressources ? *(Plusieurs réponses possibles)*

- HTTP remplacé par RDF
- URL généralisés aux URI ou IRI
- HTML remplacé par XSL
- HTML remplacé par RDF
- HTML remplacé par CSS

Vérifier

Enregistrer

Afficher la réponse

Vous avez utilisé 0 essais sur 3

HISTORIQUE DES SOUMISSIONS

INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

quizz de suivi entre les vidéos

Le site Web de la BBC est structuré et augmenté par le Web de Données. Le **catégorisation** des animaux pour organiser les documentaires sur la vie sauvage est directement issue de requêtes au web de données.

## Distribution

The Great white shark can be found in a number of locations including: **Africa, Australia, Europe, Mediterranean, North America, South America**. Find out more about these places and what else lives there.

## Great white shark size



## Classification

- Life
- Animals
- Vertebrates
- Cartilaginous fish
- Sharks
- Mackerel sharks
- Lamnidae
- Carcharodon
- Great white shark

INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

A screenshot of a web application interface. At the top, there is a phylogenetic tree diagram showing the evolutionary relationships between various groups of animals, including sharks, rays, bony fishes, amphibians, reptiles, birds, and mammals. Below the tree, there is a section titled "Cartilaginous fish" with a "Show All 14 Cartilaginous fish" button. Underneath, there are several small thumbnail images representing different types of cartilaginous fish, such as sharks and rays. The interface also includes a video player at the bottom with a progress bar and a volume control icon.

démos d'applications

### RETROUVER LES ÉTAPES DE STANDARDISATION

Voici les étapes de standardisation au W3C dont on a enlevé des mots clefs remplacés par AAAA, BBBB, CCCC, DDDD et EEEE.

- AAAA
- BBBB recommandation
- CCCC recommandation
- last DDDD
- working EEEE

INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

#### EXERCICE (5 points possibles)

Indiquez le mot clef anglais remplacé par AAAA.

Indiquez le mot clef anglais remplacé par BBBB.

Indiquez le mot clef anglais remplacé par CCCC.

Indiquez le mot clef anglais remplacé par DDDD.

Indiquez le mot clef anglais remplacé par EEEE.

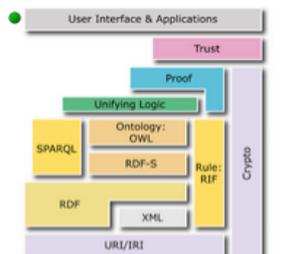
NB : Certaines des corrections s'afficheront avec des expressions régulières `ls*` entourant la bonne réponse : ne tenir compte évidemment que du terme qui se trouve entre les 2 expressions régulières !

Vérifier Enregistrer Afficher la réponse Vous avez utilisé 0 essais sur 2

HISTORIQUE DES SOUMISSIONS INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

#### PILE DU WEB SÉMANTIQUE (1 point possible)

Cliquez sur le nom du modèle de base des descriptions de ressources sur le web de données



Pile des standards du Web de données W3C®

Vérifier Enregistrer Afficher la réponse Vous avez utilisé 0 essais sur 2

HISTORIQUE DES SOUMISSIONS INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

### 5 ÉTOILES (5 points possibles)

Retrouvez les cinq étapes de la publication de données ouvertes et liées sur le Web de qualité 5 étoiles. On vous redonne le document fondateur.

\* les données doivent être...

\*\* les données doivent être...

\*\*\* les données doivent être...

\*\*\*\* les données doivent être...

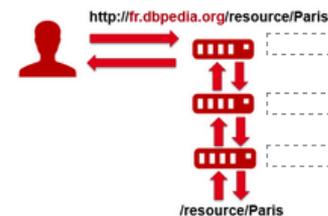
\*\*\*\*\* les données doivent être...

Vérifier Enregistrer Afficher la réponse Vous avez utilisé 0 essais sur 2

HISTORIQUE DES SOUMISSIONS INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

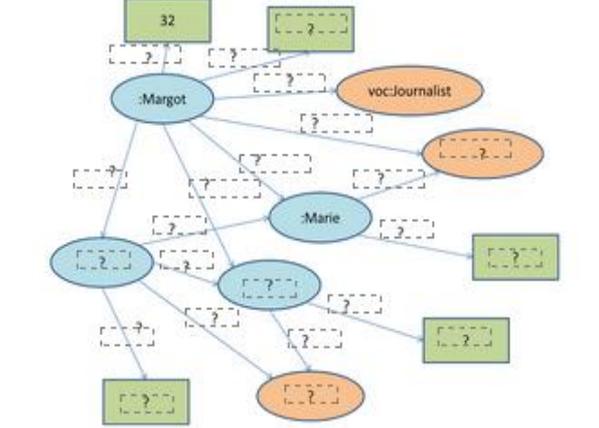
#### ROUTAGE D'UN APPEL (1 point possible)

Reconstituez l'ordre dans lequel un appel à cet URI HTTP passera par les machines gérant les différents sous-domaines.






Vérifier Enregistrer Afficher la réponse Vous avez utilisé 0 essais sur 2



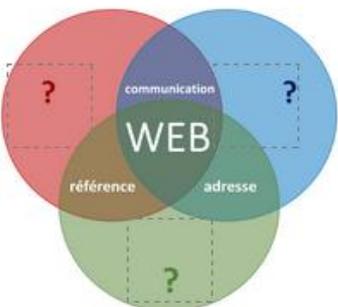






#### ARCHITECTURE WEB (1 point possible)

Recréez le diagramme du cours en plaçant les labels sur les points d'interrogation.






Glissez-déposez les trois labels pour refaire le diagramme du cours.

Vérifier Enregistrer Afficher la réponse Vous avez utilisé 0 essais sur 2

HISTORIQUE DES SOUMISSIONS INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

# exercices de chaque semaine

## LA COMMANDE CURL

Pour terminer nous allons accéder au web de données non plus par un navigateur mais par des librairies permettant de faire des appels HTTP comme le ferait un logiciel.

Nous utiliserons l'utilitaire **CURL** (<http://curl.haxx.se/>)

Sous Mac OS, l'utilitaire est probablement déjà présent. Vous pouvez le tester en demandant son manuel dans une console (fenêtre de commande shell):

```
home -- bash -- 77x38
Last login: Wed Jan 28 20:29:42 on console
liviyo:~$ curl
curl: try 'curl --help' or 'curl --manual' for more information
liviyo:~$ curl --help
Usage: curl [options...] <url>
Options: (H) means HTTP/HTTPS only, (F) means FTP only
  --anyauth      Pick "any" authentication method (H)
  -a, --append  Append to target file when uploading (F/S/FTP)
  --basic       Use HTTP Basic Authentication (H)
```

Sous d'autres plateformes vous devrez peut-être installer celui-ci.

**Avant tout:** essayez la commande "curl" dans une fenêtre de commande sur votre machine pour voir si elle est déjà installée.

**Notez** que le site fournit un **assistant d'installation pour CURL** vous guidant dans votre choix parmi les versions convenant aux différents systèmes d'exploitations. Après quoi, ouvrez une fenêtre de commande et placez vous dans le répertoire où vous avez installé le logiciel.

Vous pouvez maintenant récupérer des fichiers de données sur le Web de données.

Par exemple pour Paris:

- Les données en HTML sont obtenues par défaut avec la commande:

```
curl -o Paris.html -L http://dbpedia.org/resource/Paris
```

Les données en RDF XML sont obtenues avec un paramètre qui ajoute une entête (header) à la requête HTTP afin d'indiquer au serveur interrogé le format voulu pour la réponse (ici RDF XML):

```
curl -o Paris-rdf+xml.txt -L -H "Accept: application/rdf+xml" http://dbpedia.org/resource/Paris
```

Voici une copie d'écran de cette manipulation sous Windows:

```
C:\>cd C:\Users\ngandon\Desktop\noooc
C:\Users\ngandon\Desktop\noooc>dir
Le volume dans le lecteur C n'a pas de nom.
Le numéro de série du volume est 38C8-8F3E

Répertoire de C:\Users\ngandon\Desktop\noooc
19/01/2015  16:00    <DIR>          .
19/01/2015  16:00    <DIR>          ..
09/11/2014  14:47                558 912 curl.exe
                2 864x2  83 752 914 944 octets 1libres

C:\Users\ngandon\Desktop\noooc>curl -o Paris.html -L http://dbpedia.org/resource/Paris
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload  Total   Spent    Left   Speed
100 1326k  100 1326k    0     0  3143k    0 --:--:-- --:--:-- --:--:-- 3143k

C:\Users\ngandon\Desktop\noooc>curl -o Paris-rdf+xml.txt -L -H "Accept: application/rdf+xml" http://dbpedia.org/resource/Paris
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload  Total   Spent    Left   Speed
100 1398k  100 1398k    0     0  3416k    0 --:--:-- --:--:-- --:--:-- 3416k

C:\Users\ngandon\Desktop\noooc>
```

Voici une copie d'écran de cette même manipulation sous Mac OS:

```
home -- bash -- 109x37
liviyo:~$ curl -o Paris.html -L http://dbpedia.org/resource/Paris
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload  Total   Spent    Left   Speed
100 1326k  100 1326k    0     0  332k    0  0:00:03  0:00:03 --:--:-- 502k
liviyo:~$ curl -o Paris-rdf+xml.html -L -H "Accept: application/rdf+xml" http://dbpedia.org/resource/Paris
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload  Total   Spent    Left   Speed
100 1398k  100 1398k    0     0  242k    0  0:00:05  0:00:05 --:--:-- 308k
liviyo:~$
```

## DE LA SYNTAXE RDF/XML À TURTLE ET RETOUR

Vous allez vérifier que les énoncés solutions des deux exercices précédents sont bien équivalents. Pour cela, vous allez utiliser l'outil **RDF Translator**, accessible en ligne à l'adresse : <http://rdf-translator.appspot.com/>

Vous utiliserez ensuite le service de validation du W3C, accessible à l'adresse <http://www.w3.org/RDF/Validator/> pour visualiser la liste de triplets produits et le graphe qu'ils composent.

INFO DE DÉBOGAGE POUR L'ÉQUIPE PÉDAGOGIQUE

QUESTIONS (6 points possibles)

### RDF TRANSLATOR

Copiez le code RDF/XML solution de l'exercice 3, qui formalise l'énoncé:

**"Margot est une femme journaliste, âgée de 32 ans, mariée à Arthur qui est un homme avec qui elle a deux enfants, Marie qui est une femme et Simon qui est un homme." Pour chaque personne on spécifie aussi explicitement le nom.**

(Nous vous rappelons le vocabulaire utilisé ici: `voc:name`, `voc:age`, `voc:hasChild`, `voc:hasSpouse`)

Collez-le dans l'onglet « Input Field ».

Indiquez bien RDF/XML en Input et demandez N3 (Turtie) comme Output. Si vous avez correctement effectué cette opération, l'outil parse le code RDF/XML. Il est valide et est donc traduit dans la syntaxe N3.

Vous allez maintenant générer à nouveau du code RDF/XML à partir du code N3 produit par RDF Translator. Pour cela, copiez le code N3 produit par RDF Translator (bouton « Copy to Clipboard ») et donnez-le lui maintenant en Input.

**ATTENTION:** Indiquez bien N3 en Input et RDF/XML en Output. **Ne pas utiliser "Pretty RDF/XML"**



Combien de lignes de code sont produites ?

Constatez qu'il y a des différences syntaxiques entre le code RDF/XML produit et le code RDF/XML original.

Quel est le nom avec son préfixe de l'élément XML qui introduit l'URI de Margot comme sujet de triplet dans le code produit par RDF translator ?

Quel est le nom avec son préfixe de l'élément XML qui introduit l'URI de Margot comme sujet de triplet dans le code original ?

Quel est le nom avec son préfixe de l'élément XML introduit par RDF Translator le fait que Margot est une journaliste

# exercices avancés

Toutes les discussions ▾ Rechercher... 🔍

Tout montrer ▾ par activité récente ▾

- ÉPINGLE
- Obtention de attestation de suivi. 4  
ÉPINGLE
- Quiz: Inférences en SKOS : problème résolu 7  
ÉPINGLE
- Le mot "subsumption" 1  
ÉPINGLE SUIVI  
PAR : L'ÉQUIPE PÉDAGOGIQUE
- Sérialisation en XML 5  
ÉPINGLE
- URI et prédicat ? 11  
ÉPINGLE
- difficulté : installation de curl sur windows 8 3  
ÉPINGLE
- Système de notation et URI dans les réponses 6  
ÉPINGLE
- N'oubliez pas de rajouter k à la fin de la réponse 4  
ÉPINGLE
- Présentez-vous 135  
ÉPINGLE PAR : L'ÉQUIPE PÉDAGOGIQUE
- Charte du Forum 3  
ÉPINGLE PAR : L'ÉQUIPE PÉDAGOGIQUE
- SemWeb.Pro 2015 1
- "Serveur" SPARQL ? 1
- Syntaxe SPARQL 1

Ce message est visible par tous.

## URI et prédicat ?

4 votes +

question posté il y a 8 mois par **Raphalice05**

ÉPINGLE

Bonjour,

J'avoue que j'ai du mal à voir comment les prédicats vont avoir des URI pertinentes. Autant j'imagine bien une URI plus ou moins unique par ressource, autant pur les concepts attachés à un prédicat, je vais avoir une forêt de prédicat défini par chaque utilisateur (ou dans chaque base de données. Qui me dit que le prédicat auteur de l'inria a le même sens que le prédicat auteur d'un autre labo ? Ca va être la guerre des concepts). D'ailleurs si il n'a pas la même URI, ça ne sera pas la même chose). Est ce qu'on ne risque pas d'avoir une multitude de déclarations contradictoires ?

Autre question, le triplet RDF ne devrait il (pourrait il pas) pas avoir lui meme une URI ?

En rapport à: [Semaine 2 / 2 Modèle de triplets et graphes](#)

**Lully**

5 votes +

il y a 8 mois - marqué comme réponse il y a 8 mois par **Ogouchi**

ÉQUIPE PÉDAGOGIQUE



@Raphalice05 : je suis tout à fait d'accord sur le risque que vous décrivez. Sauf que je pense que l'ensemble des organismes et personnes qui se lancent dans l'utilisation des technologies du web sémantique en ont également plus ou moins conscience.

Et je vois grosso modo que les communautés d'intérêt s'organisent consensus *de facto* autour de terminologies communes.

forum

# Bilan

## Session 1



- 4000 participants, 25-35 ans, 466 finalistes ~11%
- 40% ingénieurs, 12% étudiants, 10% enseignants, 29% autres (biblio, santé, etc.)
- 36% néophytes, 37% consolidation

# Bilan

## Session 1



- 4000 participants, 25-35 ans, 466 finalistes ~11%
- 40% ingénieurs, 12% étudiants, 10% enseignants, 29% autres (biblio, santé, etc.)
- 36% néophytes, 37% consolidation
- 76% passaient plus de 2h/s et 40% plus de 3h/s
- 80% ont regardé les démos/screencasts

# MOOC rejoué

## Session 2



- Printemps 2016
- Identique à Session 1 à 90%
  - Intégration retours forums, e.g. demandes transcription
  - Analyse logs, e.g. exercices difficiles, énoncés à reformuler
- Beaucoup moins de temps nécessaire à l'animation du forum

# Bilan

## Session 2



- 5156 inscrits, 423 attestations ~8,2%
- 51% salariés privé/public, 11.45% recherche d'emploi, 10 % étudiants/chercheurs (doctorants)
- 1/3 de femmes
- 47% viennent des sciences humaines (y compris info-com) : public multidisciplinaire
- 45% débutants, 39% connaissent le sujet (souhaitent approfondir)
- Motivations variées : cadre pro, perso, projet pro (plaisir : 52.56%, travail 45.81%, nouvelles compétences 61.15%)
- 4 heures par semaine consacrées au MOOC

# MOOC rejoué et pérennisé

## Session 3



- Automne 2017, même taux de participation
- Plus de possibilité de rejouer le MOOC (besoin toujours là)
- Vidéos accessibles sur Canal U
- Ouverture prochaine *permanente* du MOOC sur FUN (sans animation)

# Complément et continuation



- 3 sessions : automne 2016, printemps 2016, automne 2017,
- ~ 2000 inscrits



EIT Digital Blended Master Data Science

# Conclusion et retour d'expérience

- Beaucoup de travail
  - pour la conception, la création du matériel pédagogique et l'animation du forum
  - le « business model » dans universités en France reste à trouver
- Réutilisation du matériel pédagogique
  - par les auteurs du MOOC à Polytech Nice Sophia et UCA
  - dans de nombreuses autres formations académiques
- Nouvelles pratiques pédagogiques
  - SI5 Polytech Nice Sophia et ailleurs : en // de cours en présentiel
  - MSc SmartEdTech de l'UCA : avec intervention ponctuelle

# Conclusion et retour d'expérience

- Enseignement académique, diffusion scientifique, formation en entreprise, formation tout au long de la vie
  - Usages très variés, différents niveaux de suivi du MOOC
  - Sollicitations pour interventions dans cadres variés
  - Besoins supérieurs aux forces
  - Besoin de ressources pédagogiques francophones

Enseignement et formation en IA en ligne  
Retour d'expérience sur le MOOC  
Web de données et Web sémantique

Catherine Faron Zucker

[faron@unice.fr](mailto:faron@unice.fr)

