

MAS-RAD – CAS-DAR

Introduction au web des données

Jérôme Garo – 08.06.2020

Sources disponibles sur [GitHub](#)

Table des matières

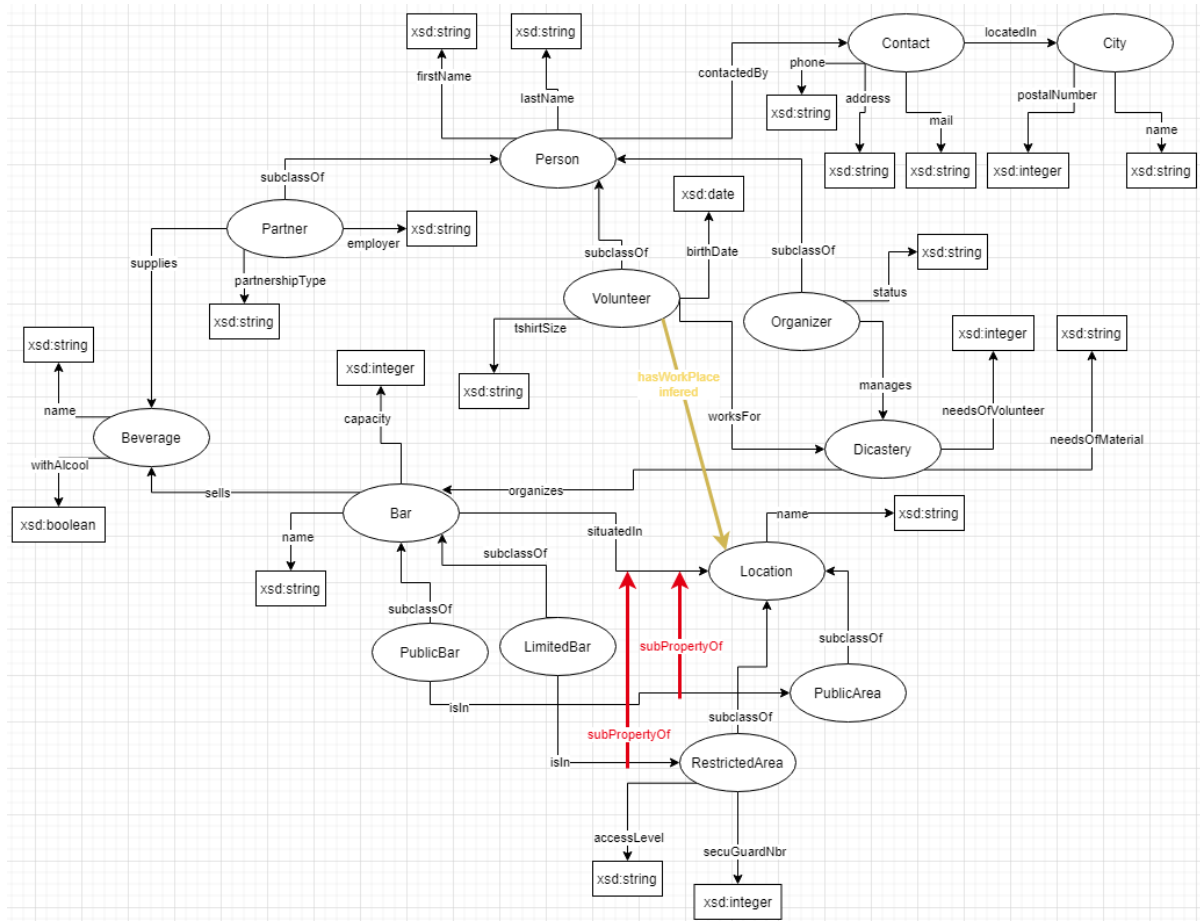
RDF schema	3
Partie 1	3
Partie 2	4
RDF datas.....	5
Personnes, contact et organisateurs.....	5
Partenaires, boissons, bénévoles et dicastères.....	6
Bars.....	7
Emplacements	7
Récupération des données.....	8
Requête 1	8
Requête 2	9
Requête 3	10
Requête 4	11
Requête 5	12
Requête 6	13
Analyse RDF.....	14
Site 1.....	14
Site 2.....	16
Site 3.....	18
Conclusions.....	19

RDF schema

Partie 1

L'ontologie présentée ici est un extrait d'une structure plus grande modélisant intégralement l'organisation d'une manifestation (festival, événement sportif, ...).

Cet extrait se concentre sur les personnes impliquées dans le fonctionnement de la manifestation ainsi que sur la partie concernant les bars.



Classes	
Nom	Description
Person	Une personne concernée par l'événement
Contact	Les informations de contact d'une personne
City	Une ville correspondant au contact
Partner	Une personne externe avec qui travaillent les organisateurs
Volunteer	Une personne travaillant bénévolement à l'organisation de l'événement
Organizer	Une personne organisant l'événement
Beverage	Une boisson qui est vendue dans l'événement
Dicastery	Un regroupement de personnes organisant une part de l'événement
Bar	Un bar...
PublicBar	Un bar accessible par tout le monde
LimitedBar	Un bar à accès limité
Location	Un endroit où se trouve un bar
PublicArea	Une zone de la manifestation accessible par tout le monde
RestrictedArea	Une zone de la manifestation accessible uniquement par certaines personnes

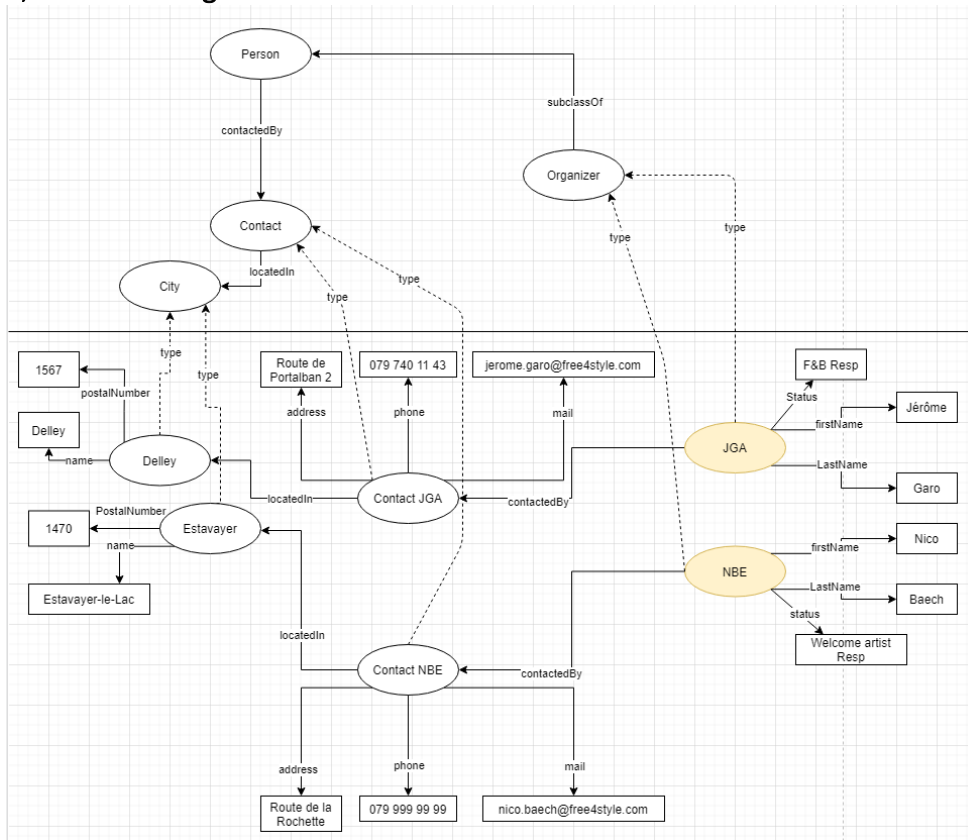
Propriétés data	
Nom	Description
firstName	Prénom d'une personne
lastName	Nom de famille d'une personne
Phone	Numéro de téléphone
Address	Rue et numéro
Mail	Adresse e-mail
postalNumber	Numéro postal d'une ville
Name	Nom générique (ville, boisson, bar endroit)
partnershipType	Le type de partenariat entre l'événement et une personne externe
Employer	L'entreprise employant une personne externe à l'événement
brthDate	La date de naissance d'un bénévole
tshirtSize	La taille de t-shirt pour un bénévole
Status	La fonction d'un organisateur (Trésorier, président, responsable de....)
needsOfVolunteer	Le nombre de bénévoles dont a besoin un dicastère
needsOfMaterial	La liste du matériel nécessaire pour un dicastère
withAlcool	La présence ou non d'alcool dans une boisson
Capacity	Le nombre de personnes maximum dans un bar
accessLevel	Le niveau d'accès pour accéder à une zone
secuGuardNbr	Le nombre d'agent de sécurité dans une zone

Propriétés objet		
Nom	Domain	Range
ContactedBy	Person	Contact
locatedIn	Contact	City
Supplies	Partner	Beverage
Sells	Bar	Beverage
worksFor	Volunteer	Dicastery
Manages	Organizer	Dicastery
Organizes	Dicastery	Bar
situatedIn	Bar	Location
isIn	PublicBar	PublicArea
	LimitedBar	RestrictedArea

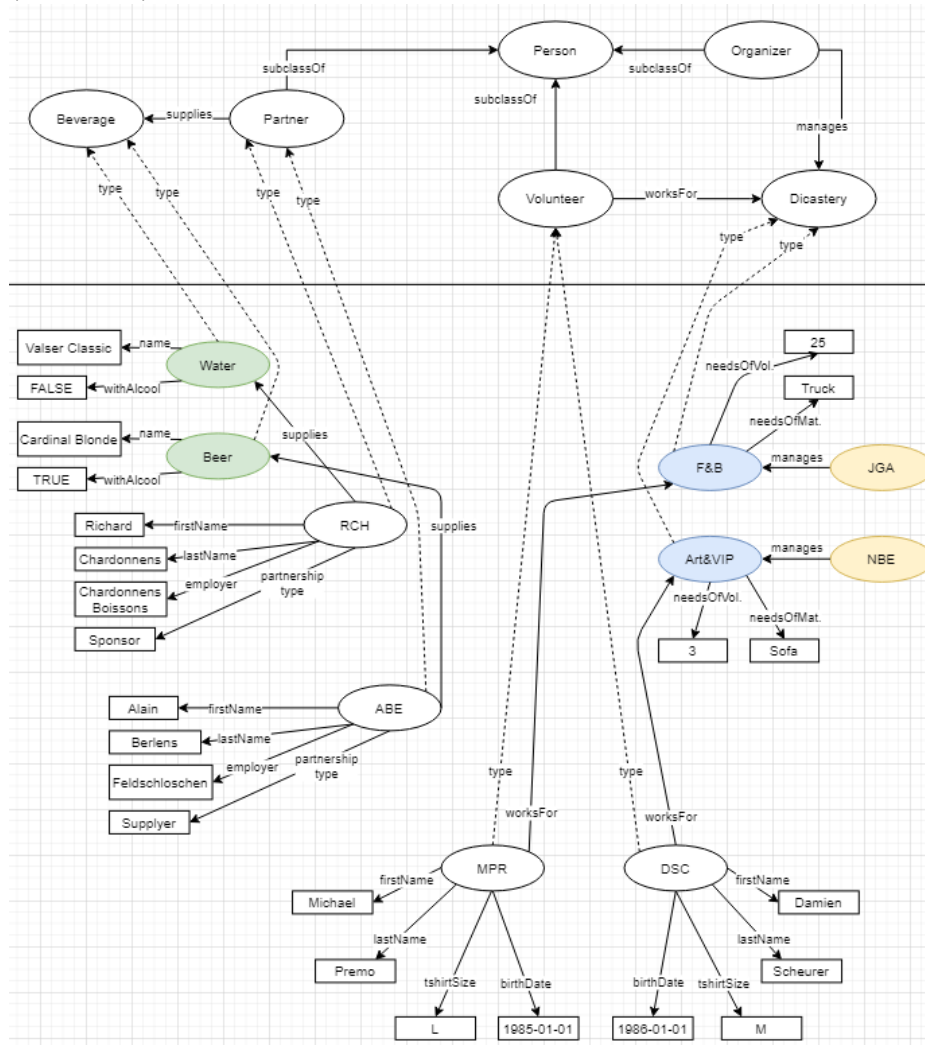
Partie 2

Une inférence "hasWorkPlace" pourrait être créée entre la classe "Volunteer" et la classe "Location" (flèche jaune sur le graph ci-dessus). Cela permettrait de ressortir rapidement l'emplacement de travail de chaque bénévole (idéal pour le bureau d'accueil), de connaître les tailles de t-shirts à distribuer à chaque emplacement de travail et de contrôler l'âge des bénévoles prévu à chaque emplacement (≥ 18 ans dans les bars).

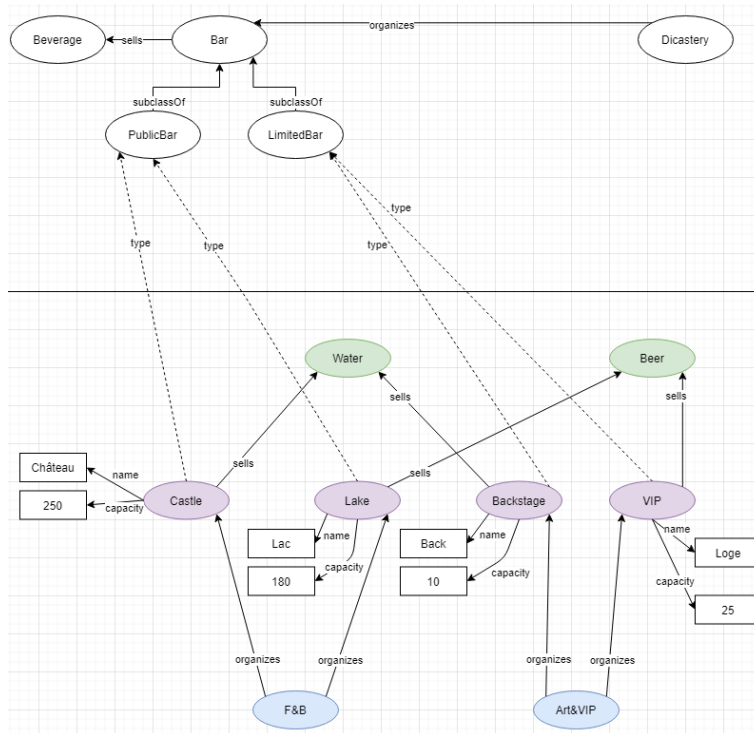
Personnes, contact et organisateurs



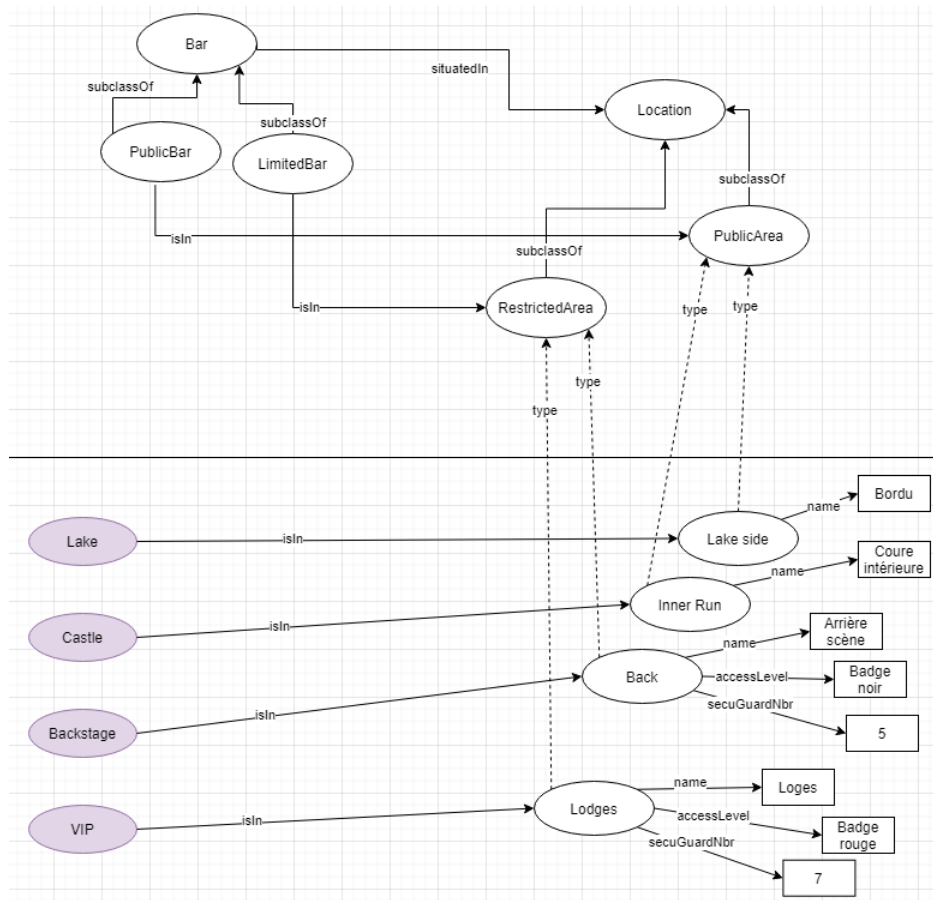
Partenaires, boissons, bénévoles et dicastères



Bars



Emplacements



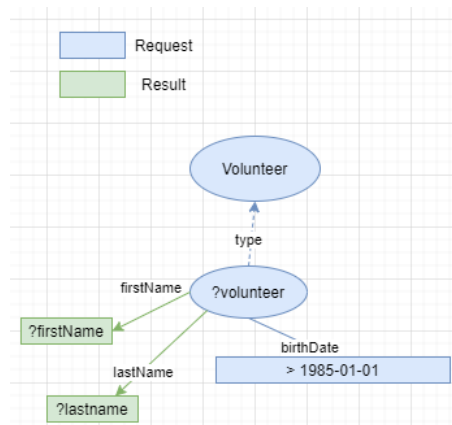
Récupération des données

Requête 1

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX project: <http://cas-dar/iwd/jga/>
SELECT ?firstName ?lastName WHERE{
    ?volunteer rdf:type project:volunteer .
    ?volunteer project:firstName ?firstName .
    ?volunteer project:lastName ?lastName .
    ?volunteer project:birthDate ?birthDate .
    FILTER(?birthDate > \"1985-01-01T00:00:00Z\"^^xsd:dateTime)}
```

Cette requête permet de retourner les noms et prénoms des bénévoles de moins de 35 ans. Cet âge a été choisi pour retourner un résultat, dans la réalité on prendrait plutôt la date qui détermine l'âge de la majorité. Grâce à cela on pourrait avoir une liste des bénévoles mineurs qui ne pourront pas avoir de tâches en contact direct avec de l'alcool.

Pour cela on recherche les sujets de type "Volunteer" et on extrait leurs noms et prénoms. Ensuite on applique un filtre sur la date désirée.



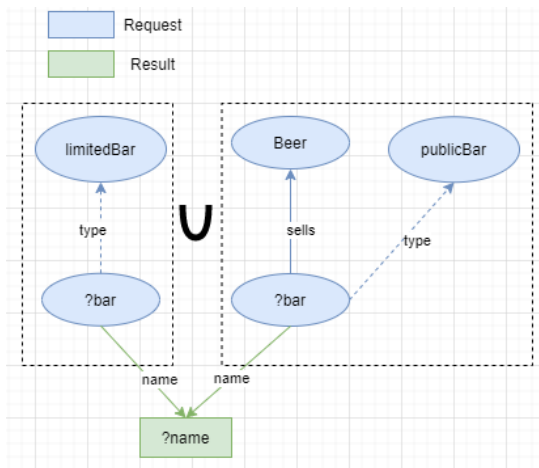
Bénévoles de moins de 35 ans :
"Damien" "Scheurer"

Requête 2

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX project: <http://cas-dar/iwd/jga/>
SELECT ?name WHERE{
    {?bar project:sells project:Beer .
     ?bar rdf:type project:publicBar}
    UNION
    {?bar rdf:type project:limitedBar}
    ?bar project:name ?name};
```

Cette requête ressort les bars qui vendent de l'alcool ou alors qui se trouvent dans une zone à accès restreint. On connaît ainsi les emplacements où on ne peut pas faire travailler les plus jeunes bénévoles (cf. Requête 1).

Pour ce faire on sélectionne les sujets de type "limitedBar" et on ajoute les sujets de type "publicBar" qui vendent de la bière. On ressort ensuite le nom des bars résultants.



Bars vendant de la bière ou en zone à accès limité :

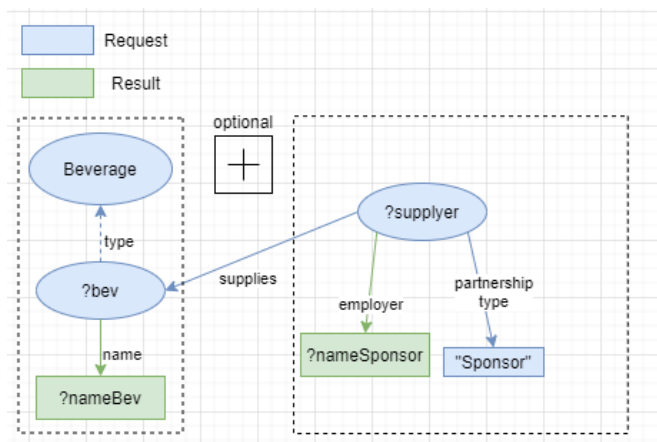
- "Lac"
- "Back"
- "Loge"

Requête 3

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX project: <http://cas-dar/iwd/jga/>
SELECT ?nameBev ?nameSponsor WHERE{
    ?bev rdf:type project:beverage .
    ?bev project:name ?nameBev .
    OPTIONAL{?supplier project:partnershipType "Sponsor" .
        ?supplier project:supplies ?bev .
        ?supplier project:employer ?nameSponsor
    }
}
```

Cette requête affiche tous les noms de boissons vendues dans le festival et y ajoute le nom du sponsor qui les fournit s'il y en a un. Cela permet de gérer la mise en avant des sponsors à proximité des points de vente où se trouvent les boissons concernées.

Pour ce faire on ressort tous les noms de boisson et on le lie au fournisseur de type sponsor concerné (s'il existe) grâce à "OPTIONAL"



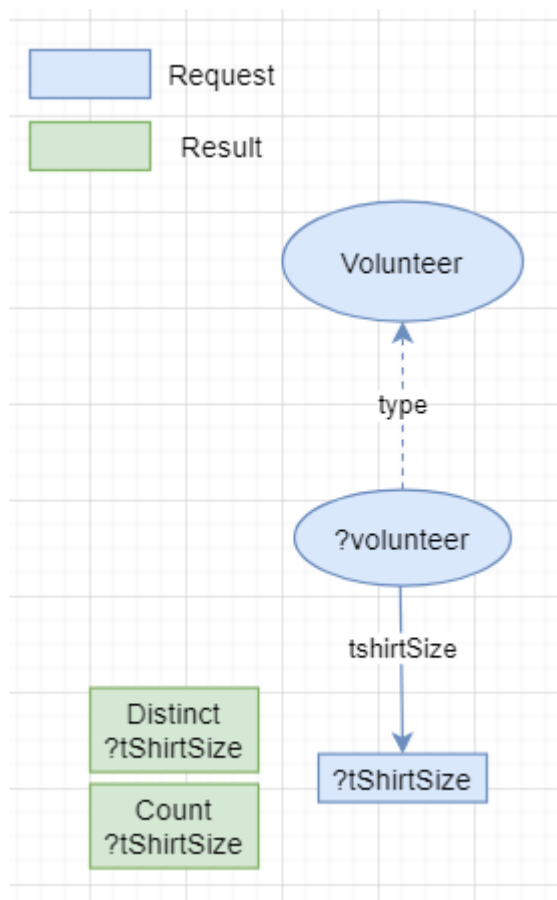
Liste des boissons et sponsors :
"Valser Classic" "Chardonnens Boissons"
"Cardinal Blonde"

Requête 4

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX project: <http://cas-dar/iwd/jga/>
SELECT DISTINCT ?tShirtSize (count (?tShirtSize) as ?numberTS) WHERE{
    ?volunteer rdf:type project:volunteer .
    ?volunteer project:tshirtSize ?tShirtSize}
"GROUP BY ?tShirtSize";
```

Cette requête affiche toutes les tailles de T-Shirts et le nombre nécessaire pour chaque taille. Cela permettra de préparer les stocks ou les commandes au fournisseur.

Pour ce faire on récupère les sujets de type "Volunteer" pour ensuite ressortir la liste des tailles de T-Shirts. On les regroupe ensuite et on les compte pour générer une colonne de total.



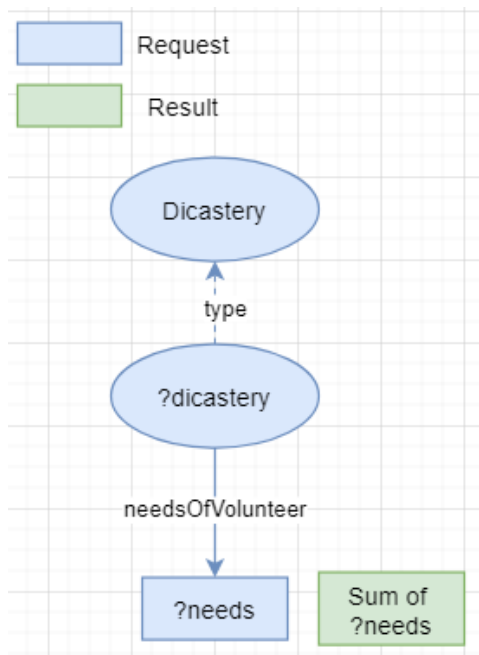
Nombre de T-Shirts :
"M" -- "1"^^<http://www.w3.org/2001/XMLSchema#integer>
"L" -- "1"^^<http://www.w3.org/2001/XMLSchema#integer>

Requête 5

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX project: <http://cas-dar/iwd/jga/>
SELECT (SUM(?needs) as ?volNeeds) WHERE{
    ?dicastery rdf:type project:dicastery .
    ?dicastery project:needsOfVolunteer ?needs
};
```

Cette requête compte le nombre total de bénévoles désirés. Cela peut être utile pour le suivi des inscriptions.

Pour ce faire on récupère les sujets de type "Dicastery" puis leurs besoins en bénévoles. On en fait ensuite la somme grâce à "SUM"



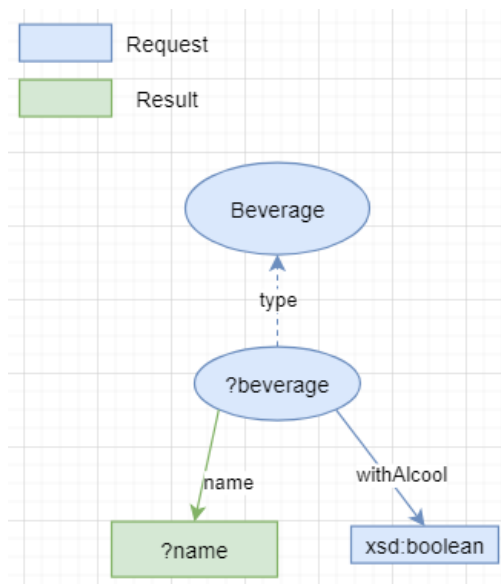
Nombre total de bénévoles désirés :
"28"^^<http://www.w3.org/2001/XMLSchema#integer>

Requête 6

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX project: <http://cas-dar/iwd/jga/>
SELECT ?name WHERE{
  ?beverage rdf:type project:beverage .
  ?beverage project:withAlcool
    \"true\"^^<http://www.w3.org/2001/XMLSchema#boolean> .
  ?beverage project:name ?name
}
```

Cette requête permet de ressortir toutes les boissons contenant de l'alcool vendue dans la manifestation. Cela peut être utile pour un futur calcul de TVA ou autre opération financière.

Pour ce faire on ressort les sujets de type boissons qui contiennent de l'alcool et on en extrait le nom.



Liste des boissons avec alcool :
"Cardinal Blonde"

Analyse RDF

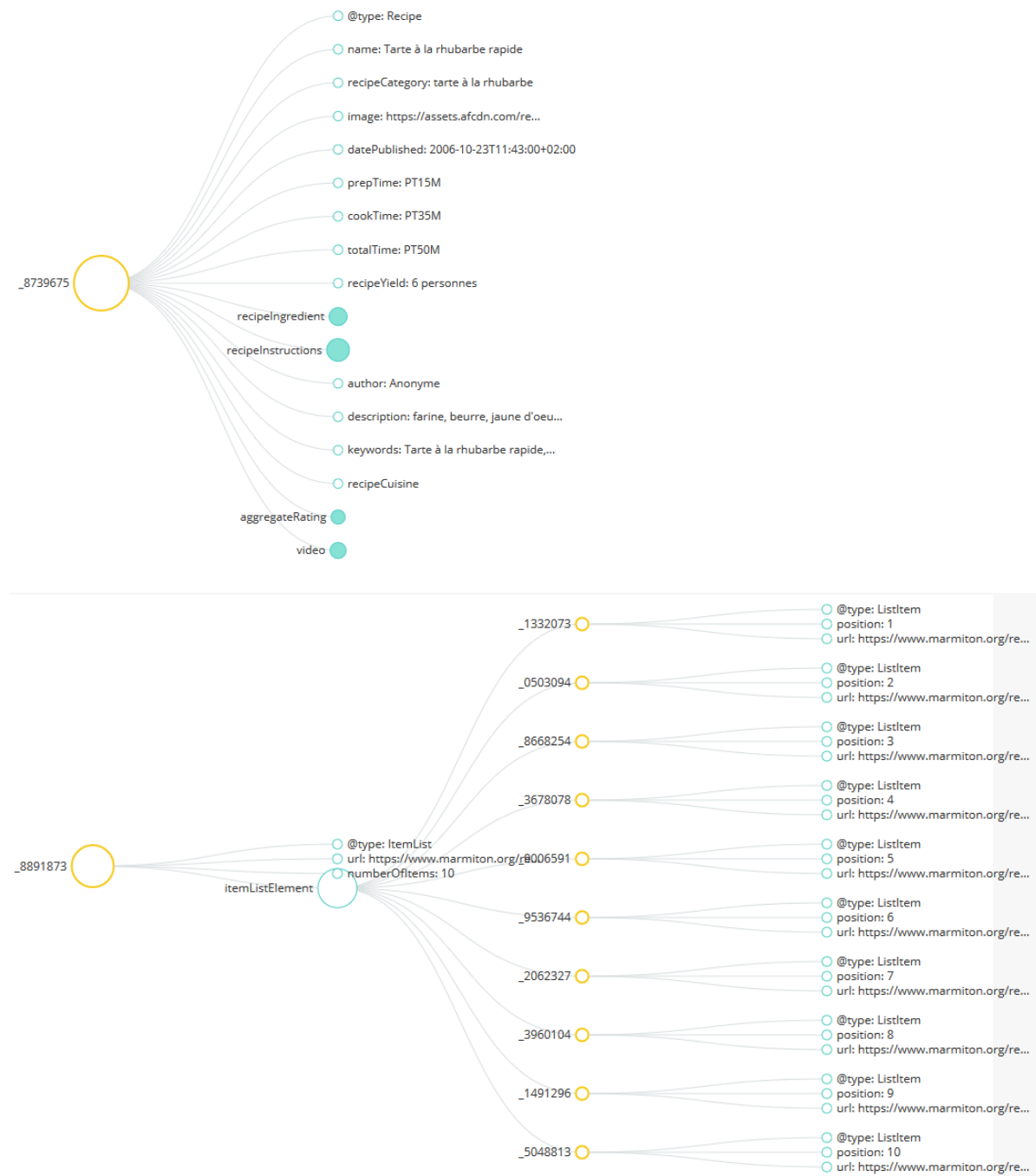
Site 1

Site : https://www.marmiton.org/recettes/recette_tarte-a-la-rhubarbe-rapide_22715.aspx

Ce site de cuisine utilise des structures JSON-LD dans ses présentations de recettes de cuisine. Sur cette page, on peut voir l'utilisation du graphe "Recipe" et "ItemList" de "shema.org"

```
{ "@context": "http://schema.org", "@type": "Recipe",  
  { "@context": "http://schema.org", "@type": "ItemList"
```

Lors de l'extraction des données voici les graphs qui en ressortent :



Tarte à la rhubarbe facile et rapide : Recette de Tarte à la ...



1 juin 2020 - Couper la **rhubarbe** (sans la peler) en petits cubes et l'étaler sur la pâte. Mélanger ensemble les 2 oeufs, le sucre et la crème + la cannelle (...

★★★★★ Note : 4,7 - 228 avis - 45 min

[Tarte à la rhubarbe](#) · [Tarte à la rhubarbe renversée](#) · [Tarte à la rhubarbe de Jeanne](#)

Tarte à la rhubarbe



Une tarte à la rhubarbe est une tarte garnie de rhubarbe. C'est un mets populaire au Royaume-Uni, où la rhubarbe se cultive depuis le XVII^e siècle, et où ses tiges se consomment depuis le XVIII^e siècle. [Wikipédia](#)

Ingrédients principaux : [Fraisier cultivé](#), [Rhubarbe des jardins](#)

Google ressort quelques informations importantes comme la durée, les ingrédients principaux, etc.

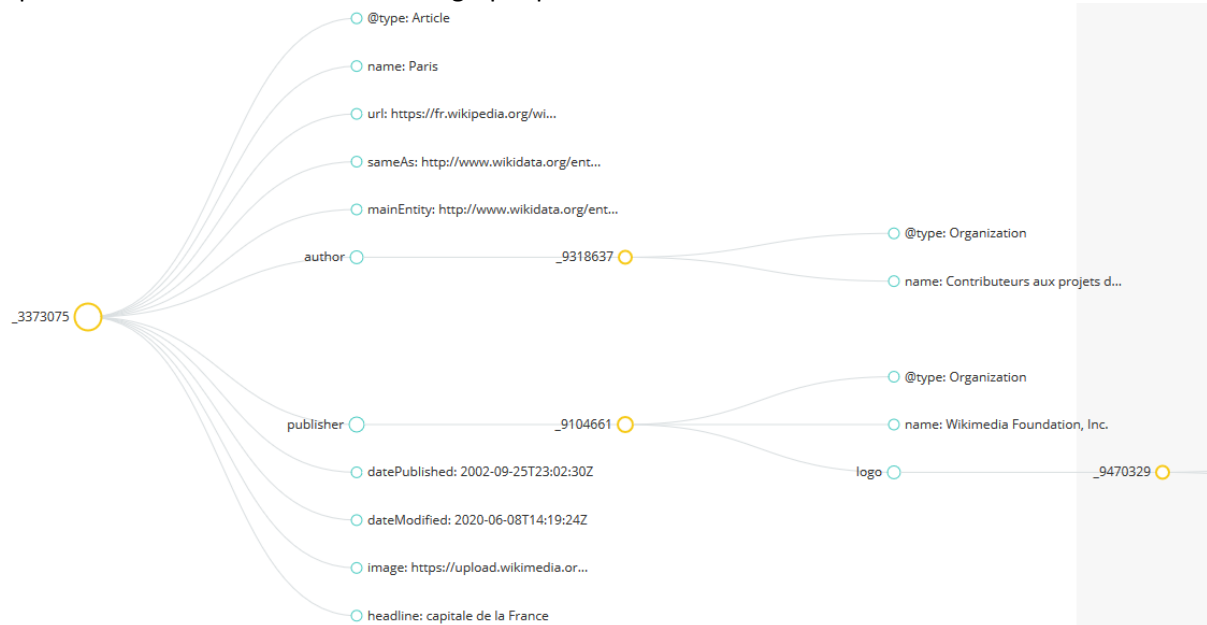
Site 2

Site : <https://fr.wikipedia.org/wiki/Paris>

Wikipedia utilise des structures JSON-LD de type "Article" pour structurer son contenu :

```
{"@context": "https://schema.org", "@type": "Article"}
```

Après extraction des données voici le graph qui en ressort :



Lors d'une recherche, Google en extrait quelques informations clés comme la population, la densité de la population, la superficie, ... Il en ressort également un extrait du texte de présentation.

[fr.wikipedia.org](https://fr.wikipedia.org/wiki/Paris) > [wiki](#) > [Paris](#) ▼

Paris — Wikipédia

Paris [pa.ʁi] est la ville la plus peuplée et la capitale de la France. Elle se situe au cœur d'un vaste bassin sédimentaire aux sols fertiles et au climat tempéré, ...

Population municipale: 2 187 526 hab. (2017) **Densité**: 20 755 hab./km²

Superficie: 105,40 km²

Région: Île-de-France (préfecture)

[Seine à Paris](#) · [Préfet de Paris](#) · [Le Bristol Paris](#) · [Géographie de Paris](#)



Paris

Capitale de la France

Paris, capitale de la France, est une grande ville européenne et un centre mondial de l'art, de la mode, de la gastronomie et de la culture. Son paysage urbain du XIXe siècle est traversé par de larges boulevards et la Seine. Outre les monuments comme la tour Eiffel et la cathédrale gothique Notre-Dame du XIIe siècle, la ville est réputée pour ses cafés et ses boutiques de luxe bordant la rue du Faubourg-Saint-Honoré.

Météo : 18 °C, vent N à 14 km/h, 58% d'humidité

Population : 2,148 millions (1 janv. 2020)

Arrondissement : Chef-lieu de 20 arrondissements

Code postal : Selon l'arrondissement, de 75001 à 75020 et 75116

Coordonnées : 48° 51' 24" nord, 2° 21' 07" est

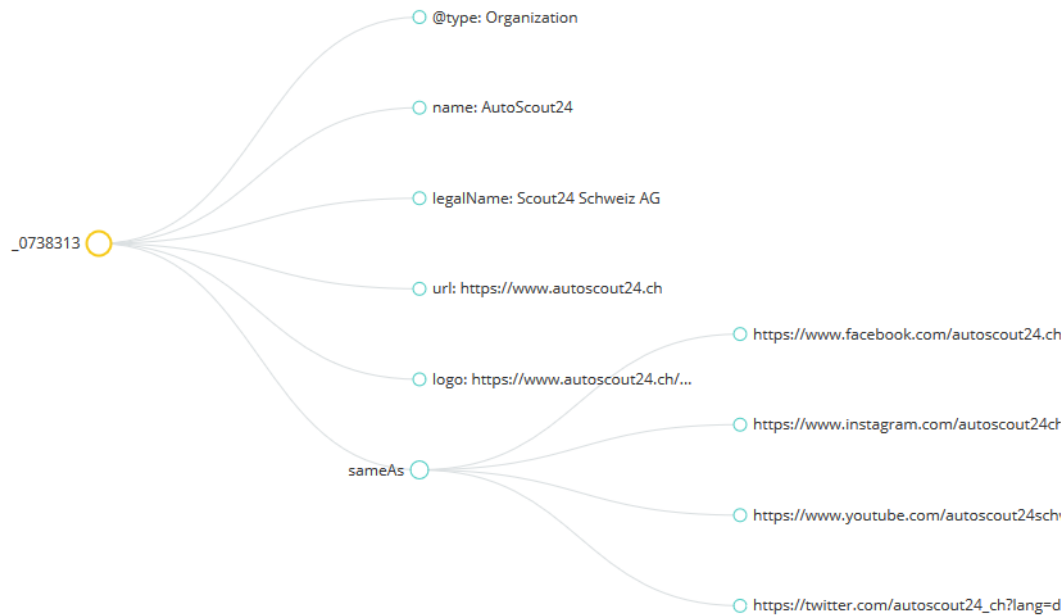
Population aire urbaine : 12 568 755 hab. (2016)

Heure locale : lundi 17:44

Site 3

Site : <https://www.autoscout24.ch/fr>

Le site Autoscout24 utilise une structure JSON-LD pour présenter leur entreprise grâce au schéma "Organization".



Dans ce cas-ci, Google ne semble pas extraire d'informations spécifiques afin de les ajouter aux résultats de recherche.

www.autoscout24.ch > ... ▼

AutoScout24: Voiture d'occasion et neuve à acheter & à ...

AutoScout24 est le marché en ligne le plus étendu & le plus connu pour les voitures en Suisse. Trouvez ici votre prochaine voiture neuve ou d'occasion.

[Recherche avancée](#) · [Véhicule utilitaire](#) · [Aperçu marques](#) · [Mon AutoScout24Se ...](#)

Conclusions

Ce projet montre l'ensemble des aspects du web sémantique, de la conception d'une ontologie à l'exploitation de ses données en passant par son implémentation. En couvrant l'ensemble de ces points on remarque très vite la grande quantité de travail que représente une telle tâche et le temps extrêmement conséquent que cela représente.

Toutefois il apparaît assez intéressant de développer sa propre ontologie lorsqu'on se trouve confronté à un sujet qui n'a pas ou très peu été couvert par celles que l'on trouve sur internet. Il est, à mon avis, plus simple de tout créer de zéro plutôt que d'essayer de mélanger plusieurs autres ontologies parfois très mal documentées.

Pour ce qui est du cours et de la matière en elle-même, je trouve très intéressant de comprendre les structures qui se cachent derrière l'organisation de l'information sur le web. Je suis cependant très étonné que, bien qu'étant une recommandation du W3C, ce sujet soit si peu traité sur internet. En effet, il est très compliqué de trouver de la documentation précise et récente (2015 est le "moins vieux" trouvé) ou même des exemples bien développés.

Grâce à ce cours nous avons vu que d'autres modèles de structuration de données sont bien plus souvent utilisés tels que les structures proposées par Apple, Facebook ou encore twitter. Il pourrait être intéressant d'étudier également ces types d'organisations puisqu'elles semblent supplanter les ontologies classiques.