L’un des format qui existe acutelement est linked open data. C’est un ensemble d’information qui est structuré de manière à être ouvert (données disponible dans un format utilisable par des ordi) et relié. Le model de la donnée est qu’on a 2 point relié par une relation et de cette manière la on peut avancer dans le graphe et voir la donnée comme un grand graphe. Il existe ainsi un langage qui permet de parcourir le graphe et de l’interoger grâce au requête sparql.

la première chose qu’on veut faire avec des données c’est les structurer et les rendre explicite pour l’ordinateur. C’est-à-dire on tape le nom de l’objet qu’on chercher qu’est une uri et on obtinet un resultat concernant cet objet.

Dbpedia, est une organisation qui construit un certain nombre de données de divers domaines, et constitue donc un énorme graphe relié à d’autre graphes grâce à des relations.

Chaque nœud d’un graphe est une url, si on tape cette url dans le navigateur on obtient l’information.

Dans le linkeded open data une sous categorie de donnée luinguistique llod qui nous interesse , parmi ces donnéées linguistique on trouve Dbnary qui fera objet de notre projet.

Dans dbnary 21 langues sont extraite, on extrait les données des wiktionary anglais almand …., wikitionaire contient une version pour chaque langue.

Pour chaque langue on a des donnée depuis 2012 jusqu’à 2020, notre travail est de faire ne sorte que sa soit mis a jour pour 2021. Notre travail est de reprendre ces shema la, il y’a 1 an une mise à jours à été faite dans la manière dont les information sont représenté notamenet les info statistique.

A la base les programme d’extractions qui prenait toutes les données extraites qui comptait le nombre d’elements et qui generais un fichier csv. C’est une methode très couteuse car on fait l’extration et puis on doit relir les donnée qui sont assez grosses puis faire une requête.

Travail : fournir un tableau de board permettant aux utilisateur exterieur de savoir quelle est la taille des informations dbnary à la fois globalalement, puis langue par langue, et puis quuelle est l’evolution de cette taille dans le temps langue par langue et globalement.

Facilement Parametrables juste en modifiant le shortcode on obtiendra un autre graphique ou bine si on rajoute des données il sera extrêmement simple de rajouter des graphiques.

Grâce aux graphique on vois un certain nombre de chose qu’on peut détecter et puis analyser et corriger s’il le faut, par exemple sur le graphiqye français on voie que tout tombe à l’eau parceque les français ont changé la manière dont les tardiuction sont representés.

C’est intéressant de travailler avec des données évolutive et collaborative par exemple en 2012 on a eu 405000 traductions de la langue française en 2020 on a eu 961000(vrai valeur ajouté).

*L’objectif du* Resource Description Framework *RDF est de rendre toutes ces données lisible, par l’humain et la machine, ayant une structure de graphe et en les connectant à l’aide des URI.* C’est-à-dire on tape le nom de l’objet qu’on chercher qu’est une uri et on obtinet un resultat concernant cet objet.

The Web is made up of resources. The one that interests us in our project is Wiktionary which is a