Le Web sémantique, données liées, graphes de connaissances

Phil Barker, Cetis LLP et Credential Engine

Big Data pour les informations sur le marché du travail
Programme de développement des capacités en ligne

16 novembre 2022

Nous encourageons la réutilisation et le mélange de cette ressource en indiquant la source.

Thèmes

- La diversité doit être accueillie favorablement.
- L'interopérabilité est une nécessité.
- Les choix techniques ont des conséquences.

Ce sont des thèmes fondamentaux qui doivent être esquissés à grands traits, mais dont les détails sont importants.

Contexte : la diversité et l'exigence de l'interopérabilité

Aperçu d'un marché du travail

Photo: CC BY-SA 3.0, by Vgrigas from Wikimedia
File:Pam Robertson, Grade 12A Teacher at Sinenjongo High School, Joe Slovo Park, Cape Town, South Africa-3320.jpg

Diversité

Acteur	Intérêts	Interagir avec
Employeurs	Trouver les bonnes personnes pour les postes, recrutement, progrès, paie, ...	Employés, candidats, autres employeurs, organisations professionnelles, gouvernement, …
Organisations dans le domaine de l’enseignement	Conception de programmes, gestion de cours, évaluation, certification, …	Étudiants, candidats, employeurs, gouvernement, organismes d'assurance qualité, …
Organisations professionnelles	Cadres pour rôles, plans de carrière, …	Particuliers, employeurs, établissements d’enseignement et d'assurance qualité, …
Organisations dans le domaine de la qualité	Certification de la conception et de l'évaluation des programmes en fonction des compétences	Organisations professionnelles et dans le domaine de l’enseignement
Gouvernement	Économie, législation, réglementation, …	Tous

Points clés

- Le marché du travail implique une multitude d'acteurs différents, chacun ayant ses propres préoccupations, pratiques et utilisations des données.
- L'interopérabilité est facilitée lorsqu'une norme commune est utilisée.
- Des normes communes sont possibles lorsqu'il existe des préoccupations, des pratiques et des utilisations communes des données.

Il n'existe pas (et il ne devrait pas exister) de norme commune à l'ensemble du marché du travail.

Cependant, il est nécessaire de partager les données.

Les bases techniques du Web sémantique, des données liées et des graphes de connaissances

RDF : *Resource Description Framework* (structure de la description de la ressource)

- Recommandations du W3C depuis 1999.
- Début des années 2000, axe sur le « **web sémantique »** : ontologies, représentation des connaissances et interférence logique.
- Depuis 2006, axe sur les « **données liées** » et l'interopérabilité.
- Depuis 2012, axe sur les « **graphes de connaissances** » et la fouille de données (*data mining*).

Pascal Hitzer (2021) *Review of the Semantic Web Field*. Communications of the ACM, Vol. 64 Nº 2 10.1145/3397512

RDF : *Resource Description Framework* (structure de la description de la ressource)

Sujet - Prédicat - Objet

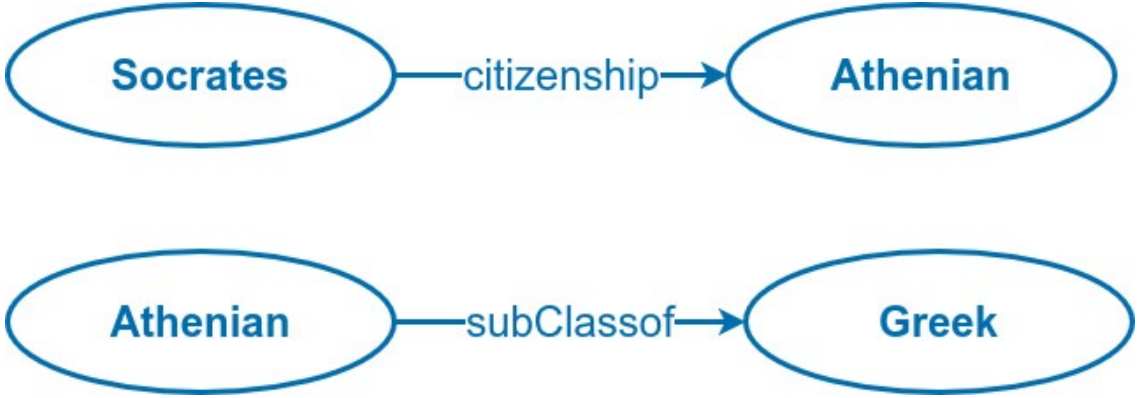
Il est possible de tout identifier avec des URI :

<https://www.wikidata.org/entity/Q913> <http://www.wikidata.org/prop/direct/P27> <https://www.wikidata.org/entity/Q844930>.

« Les gens pensent que le RDF est un problème parce que c'est compliqué. La vérité est bien pire. Le RDF est extrêmement simple et vous permet de travailler avec des données du monde réel et des problèmes qui sont vraiment très compliqués. »

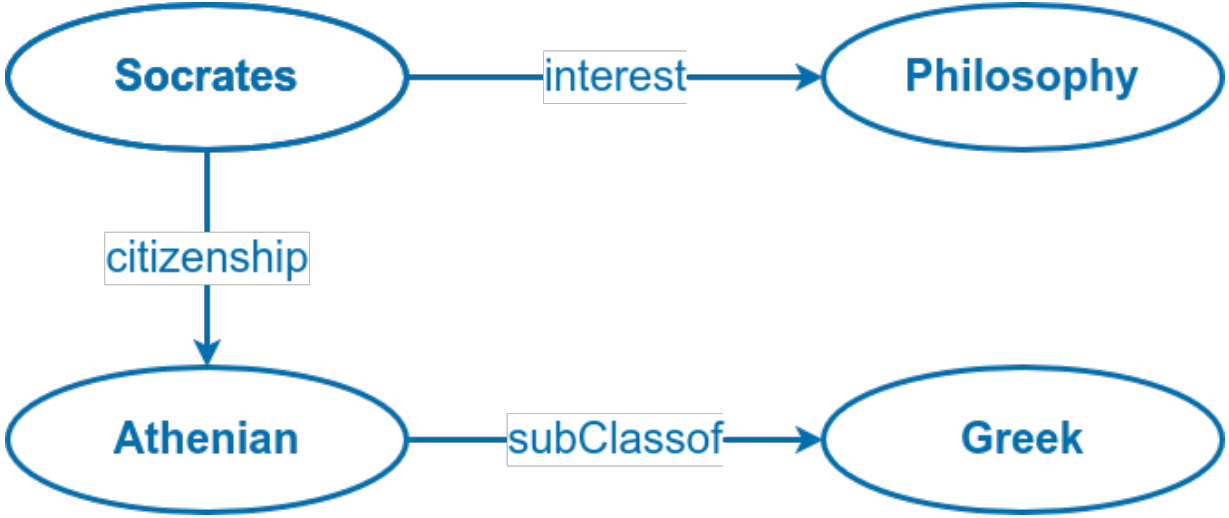
Dan Brickley, Schema.org et Google ; Libby Miller, BBC (source)

RDF et sémantique



Par conséquent : Socrate est grec.

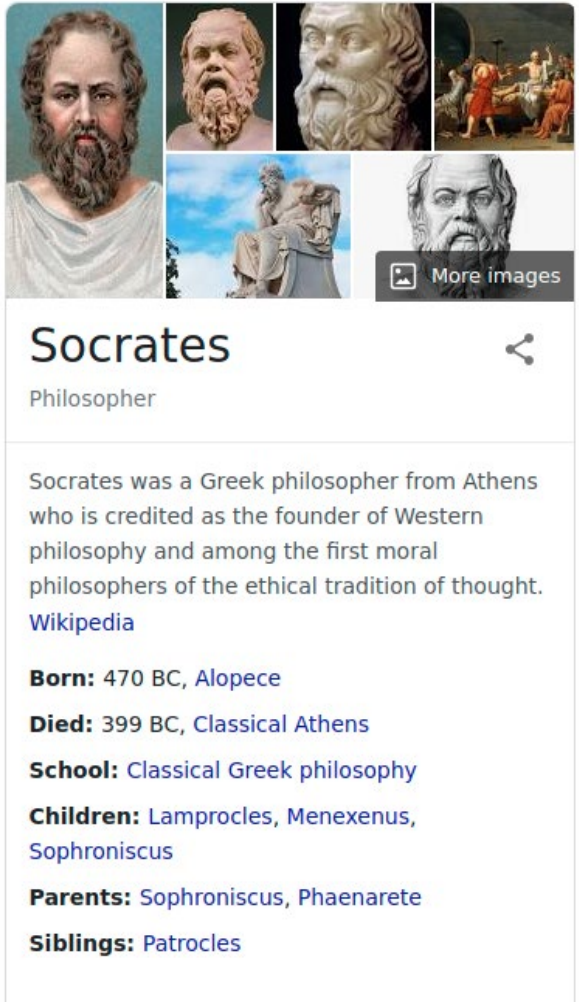
Graphes de connaissances



- Ajouter des données supplémentaires.
- L'objet d'une déclaration peut être le sujet d'une autre.
- Connecter les déclarations dans un « graphe ».
- Commencer à répondre à des questions telles que « Combien de Grecs sont intéressés par la philosophie ? »

URI et données liées

- Tout identifier avec un URI globalement unique



L'image est le « panneau de connaissances » de Google issu d'une recherche sur Socrate.

URI et données liées

- Tout identifier avec un URI globalement unique
- Permet la désambiguïsation
  - Q : Quand Socrate n'est pas grec ?
  - A : quand Socrate n'est pas grec <https://www.wikidata.org/entity/Q102331>, not Σωκράτης: <https://www.wikidata.org/entity/Q913>

Sócrates

Brazilian footballer

Sócrates Brasileiro Sampaio de Souza Vieira de Oliveira, simply known as Sócrates, was a Brazilian footballer who played as a midfielder. His medical degree and his political awareness, combined with style and quality of his play, earned him the nickname "Doctor Socrates". Wikipedia

Born: 19 February 1954, Belém, State of Pará, Brazil

Died: 4 December 2011, Albert Einstein Israelite Hospital, São Paulo, State of São Paulo, Brazil

Height: 1.92 m

Spouse: Katia Bagnarelli (m. ?–2011)

Children: Gustavo Vieira de Oliveira, MORE

Siblings: Raí

Parents: Guiomar de Oliveira, Raimundo de Oliveira

L'image est le « panneau de connaissances » de Google issu d'une recherche sur Socrate.

URI et données liées

- Tout identifier avec un URI globalement unique
- Permet la désambiguïsation
  - Q : Quand Socrate n'est pas grec ?
  - A : quand Socrate est Sócrates
- Peut suivre le lien pour plus de données
- Indépendant du langage humain

Socrates (Q913)

ancient Greek philosopher

Sokrates

In more languages

Configure

Language	Label	Description
English	Socrates	ancient Greek philosopher
British English	Socrates	classical Greek Athenian philosopher
Spanish	Sócrates	filósofo griego clásico ateniense
Greek	Σωκράτης	φιλόσοφος της Αρχαίας Ελλάδος (-470/-399)
Afrikaans	Sokrates	Griekse filosoof
Amharic	□□□□□	No description defined
Aragonese	Socrates	No description defined
Old English	Socrates	No description defined
Arabic	سقراط	أغريقي قديم
Egyptian Arabic	سوقراط	No description defined
Assamese	□□□□□□□□	No description defined
Asturian	Sócrates	No description defined
Azerbaijani	Sokrat	No description defined
South Azerbaijani	سقراط	No description defined
Bashkir	Сократ	No description defined
Bavarian	Sokrates	No description defined
Central Bikol	Socrates	No description defined

L'image est l'entrée de Wikidata pour le nom de Socrate dans plusieurs langues.

Résumé : Normes et interopérabilité

Si des organisations utilisent des données similaires pour des éléments similaires et qu'elles veulent coopérer, elles peuvent utiliser la même norme relative aux données pour partager des données (interopérabilité).

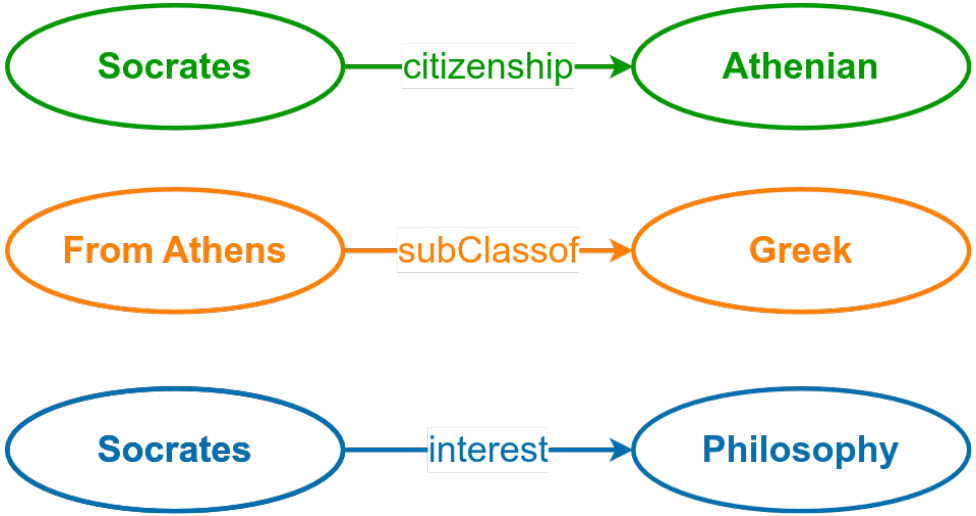
- Le département RH peut envoyer des données sur les offres d'emploi au partenaire de recrutement.
- L'éditeur peut fournir des données sur les ressources d'apprentissage à une école.

Dans chaque domaine, les systèmes informatiques mettent en œuvre les normes pertinentes.

MAIS : des normes différentes sont adaptées à des domaines différents

- Norme RH pour décrire les salaires et les compétences requises pour un emploi
- Norme LMS pour décrire les cours et les résultats d'apprentissage

RDF et harmonisation



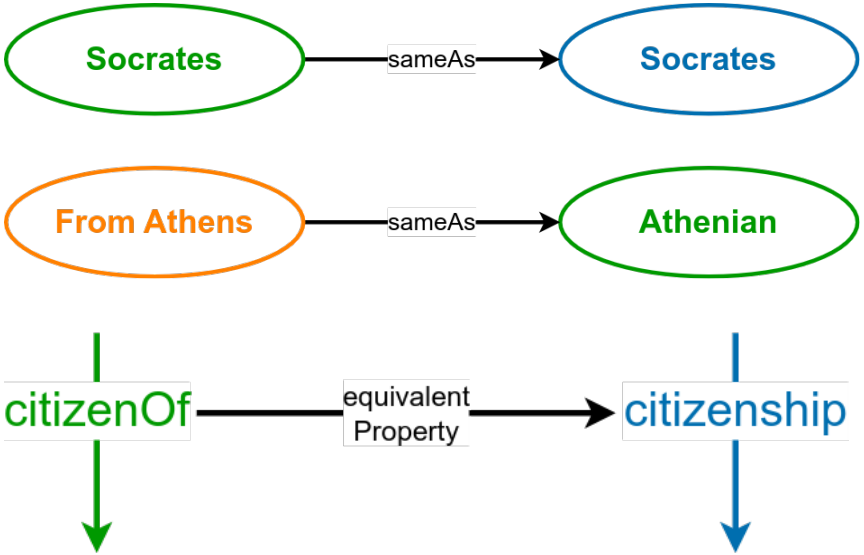
Différentes normes RDF utilisent le même **métamodèle**

(c'est-à-dire les déclarations sous forme de triplets impliquant sujet - prédicat - objet, identifiées avec les URL).

Ici, les déclarations de couleur vert, orange et bleu représentent des données provenant de trois normes RDF différentes.

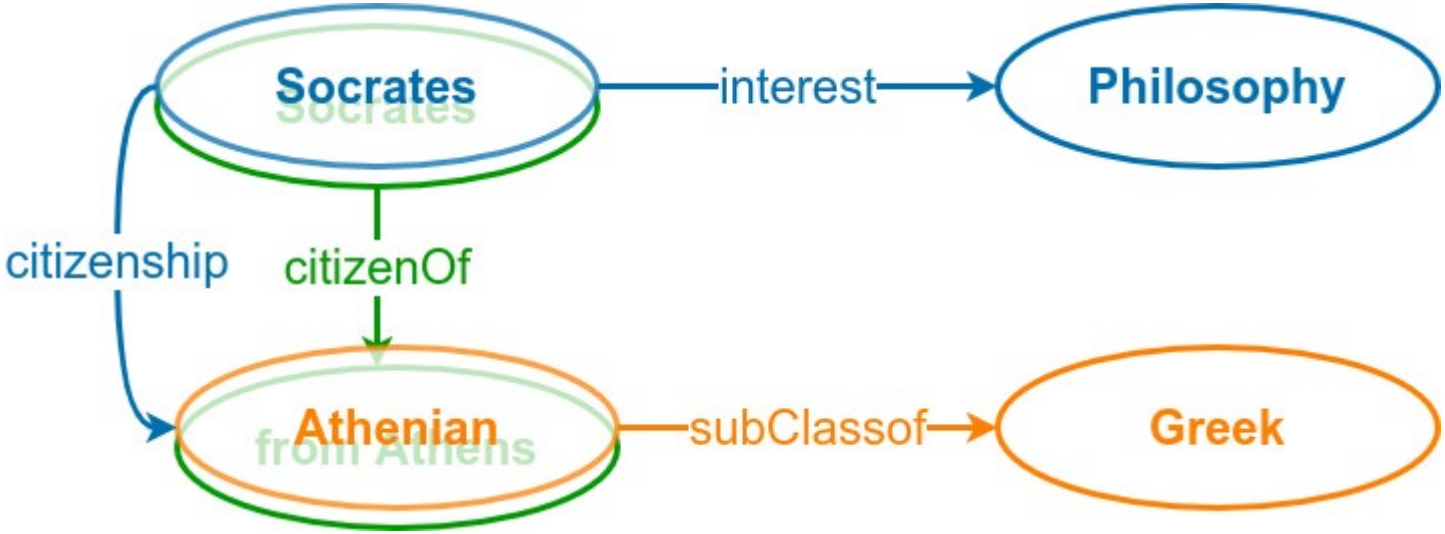
Le RDF permet de mélanger et d'établir des correspondances

Les termes de normes différentes peuvent être utilisés ensemble.



Le RDF possède des normes permettant d'affirmer l'équivalence entre les termes.

Le RDF permet de mélanger et d'établir des correspondances



- Même graphe de connaissances que précédemment, mais créé en mélangeant et en établissant une correspondance entre des termes provenant de normes RDF différentes.

Points clés

RDF :

- Cadre stable et éprouvé pour des normes ouvertes.
- Fondement du Web sémantique/données liées/graphes de connaissance.
- Permet de déduire des informations qui ne figurent pas explicitement dans les données.
- Décentralisé : lien entre les jeux de données.
- Possibilité de créer de grands jeux de données publiques.
- Harmonisation sur un métamodèle commun : interopérabilité entre des normes de différents domaines

Credential Engine : transparence des titres, CTDL et registre des titres de compétences

Credential Engine : son activité

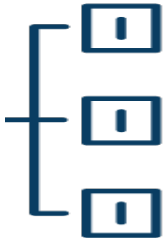
Notre mission : Credential Engine est un organisme à but non lucratif dont la mission est de cartographier le paysage des titres relatifs aux compétences à l'aide d'informations transparentes et cohérentes, tout en soutenant la création de ressources permettant aux personnes de trouver les voies qui leur conviennent le mieux.

Notre vision : Nous envisageons un avenir où des millions de personnes partout dans le monde ont accès à des informations sur les titres relatifs aux compétences leur permettant de connaître toutes les possibilités d'apprentissage, de progrès et de carrières enrichissantes.

Comment nous réalisons notre mission et notre vision :

- Construction d’une infrastructure et d’un langage communs
- Collaboration
- Autonomisation

Les technologies ouvertes de Credential Engine



Credential Transparency Description Language (CTDL)

Schéma de données ouvertes liées décrivant plus de 800 critères pour les titres, compétences, emplois et fournisseurs.

Outils de publication du CTDL

Les fournisseurs utilisent l'API ou un des outils du système de publication pour convertir les informations en CTDL, les publier dans le Registre et créer des identifiants uniques mondiaux (CTID).

Registre des titres

Le Registre collecte et relie les données relatives aux titres, aux compétences et aux emplois décrits dans la base de données CTDL sous la forme d'un graphe de données ouvertes et liées à des outils et applications puissants, dont Credential Finder.

CREDENTIAL FINDER

Credential Finder

Outil pour visualiser et explorer les informations stockées dans le Registre

Site technique : https://credreg.net

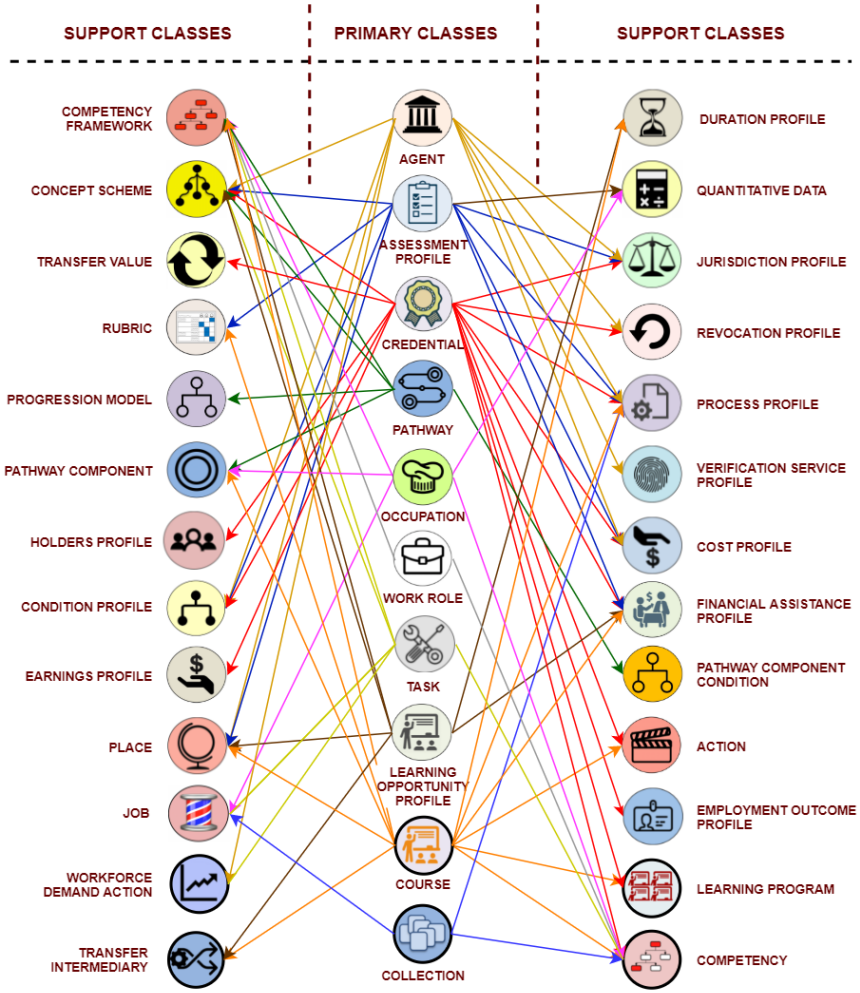
Credential Transparency Description Language (CTDL)

Un langage commun sous licence ouverte, modelisé sur les spécifications du World Wide Web Consortium (W3C) pour le Web sémantique, afin de permettre la transparence et la comparabilité des informations relatives aux titres et aux compétences.

Trois schémas :

1. CTDL : titres et classes adjacentes
2. CTDL-ASN : compétences
3. QDATA : données quantitatives

CDL Handbook (manuel du CDL) via https://credreg.net



Termes du *Credential Transparency Description Language* (CDL)

Plus de 800 termes pour les données ouvertes liées fournissant des descriptions détaillées des titres et des compé

ceterms:Certification	
URI:	https://purl.org/ctdl/terms/Certification
Shorthand URI:	ceterms:Certification
Label:	*en-US:* Certification
Definition:	*en-US:* Time-limited, revocable, renewable credential awarded by an authoritative body for demonstrating the knowledge, skills, and abilities to perform specific tasks or an occupation.
Comment:	*en-US:* Certifications can typically be revoked if not renewed, for a violation of a code of ethics (if applicable) or proven incompetence after due process. Description of revocation criteria for a specific Certification should be defined using Revocation Profile.
Type of Term:	rdfs:Class
Status:	vs:stable
Subclass Of:	ceterms:Credential ceterms:Credential
	ceterms:accreditedBy ceterms:isRequiredFor ceterms:accreditedI ceterms:administrat ceterms:advancedS ceterms:alternateN ceterms:appealProc ceterms:approvedBy ceterms:approvedIn ceterms:assessment ceterms:audienceLe

Certification : titre de compétences limité dans le temps, révocable et renouvelable, délivré par un organisme faisant autorité pour démontrer les connaissances, les compétences et les capacités à exécuter des tâches spécifiques ou une profession.

Exemple de données provenant du Credential Finder

Associate's Degree

A.A.S. in Dental Assisting

Ivy Tech Community College of Indiana

Last updated: Dec 22, 2020

Compare

About this Credential

Basic information about the Credential

The Dental Assisting program at Ivy Tech is taught by instructors who have all worked as dental assistants in the field. Students will get to study in a state-of-the-art facility learning a variety of skills involving radiography, dental materials, infection control, chairside, and general dentistry as well as dental specialties. For more information about dental assisting, visit the American Dental Assistant's Association website. The Dental Assisting program is a selective admission program. When you apply to the college, you would be accepted into the Healthcare Specialist program with a concentration in Dental Assisting while you complete the prerequisite requirements. The Dental Assisting program accepts a limited number of students each year and there is a separate application process.

View this Credential

Credential Type

Related Organization(s)

Organizations that own, offer, or perform other services related to the Credential

Owned and Offered By

Ivy Tech Community College of Indiana

Connections

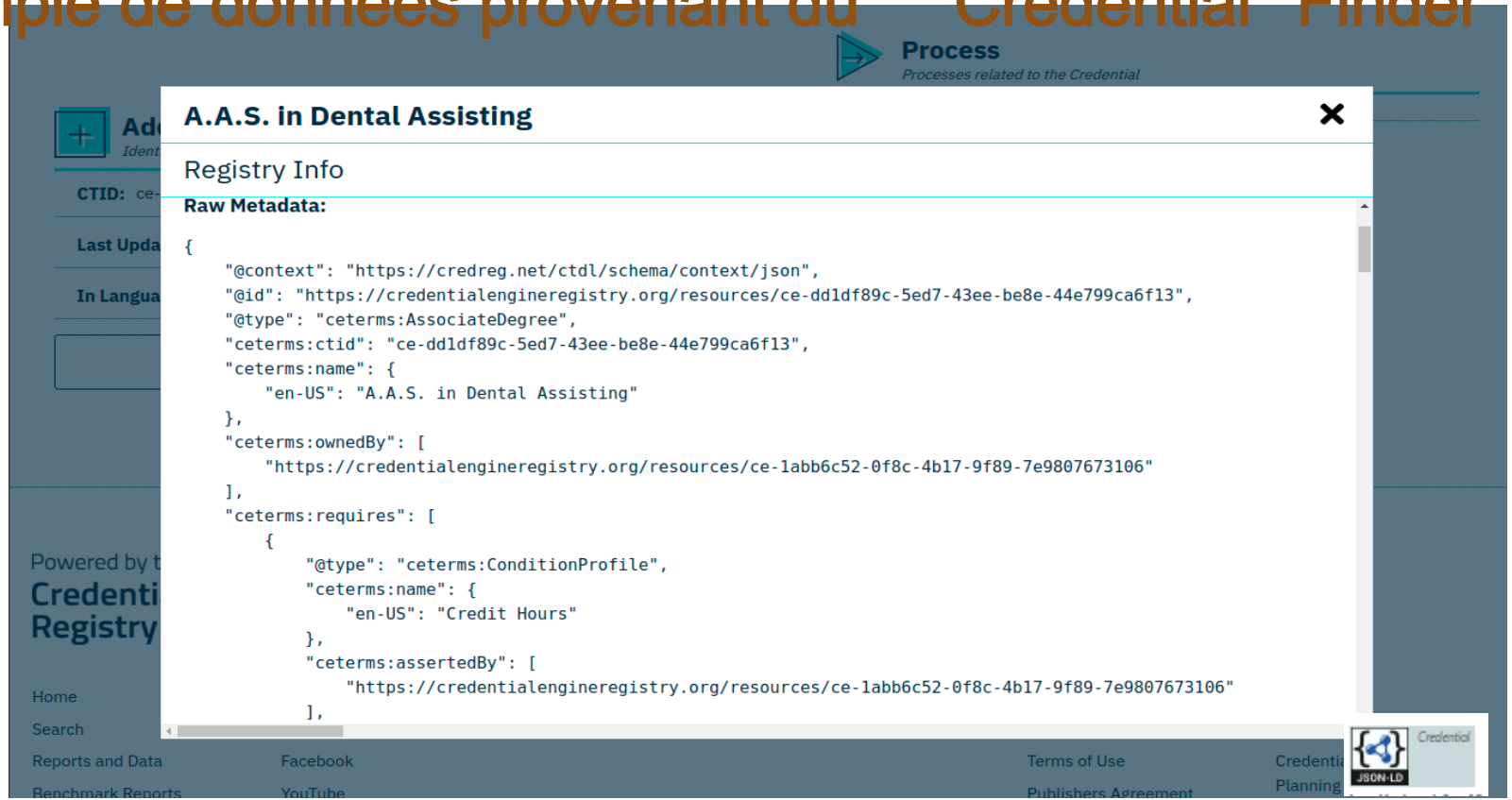
Relationships between the Credential and other data in the Registry

Related Competencies

Teaches 50 Competencies

Toutes les informations relatives à ce titre proviennent directement de credentialfinder.org

Exemple de données provenant du Credential Finder



Toutes les informations relatives à ce titre proviennent directement du Credential Registry (registre des titres) dans le format JSON-LD.

Principaux points à retenir

- Credential Engine est une organisation à but non lucratif visant à collaborer à l'amélioration des perspectives en matière d'éducation et d'employabilité.
- Le CTDL est un langage basé sur le RDF pour décrire les titres (qualifications) et les éléments connexes.
- Les données utilisant le CTDL peuvent être stockées dans des registres tels que le Credential Registry.

CTDL et justificatifs vérifiables (en bref)

Les *Verifiable Credentials* (justificatifs vérifiables) du W3C

- Norme pour les assertions évidentes inviolables des réussites d'un individu.
  - Modèle de données en RDF https://www.w3.org/TR/vc-data-model/
  - Protocoles pour l'émission de ces données dans le respect de la vie privée.
- Les exemples pédagogiques de la norme donnent très peu de détails.
- Un groupe de la communaité du W3C s'est penché sur les cas d'utilisation dans le domaine de l'éducation.
- Base de 1EdTech(*) OpenBadges V3

(* était IMS)

Exemple de VC du W3C

```
{
"@context": [
    "https://www.w3.org/2018/credentials/v1",
"https://www.w3.org/2018/credentials/examples/v1"
],
  "id": "http://example.edu/credentials/3732",
"type": ["VerifiableCredential", "UniversityDegreeCredential"],
"issuer": "https://example.edu/issuers/565049",
"issuanceDate": "2010-01-01T00:00:00Z",
"credentialSubject": {
    "id": "did:example:ebfeb1f712ebc6f1c276e12ec21",
"degree": {
      "type": "BachelorDegree",
"name": "Bachelor of Science and Arts"
}
}
}
```

Il ne suffit pas de dire que vous avez un diplôme.

Le CTDL dans un VC du W3C

```
{"@context": [...],
...
  "credentialSubject": {
    "id": "did:example:ebfeb1f712ebc6f1c276e12ec21",
"achievement":  [
      { "id": "res:ce-49247727-7cbd-4f0e-8179-e51036d1f14a", // the degree achieved
"type": "ceterms:BachelorDegree",
"ceterms:name": {"en-US": "Bachelor’s Degree"},
"ceterms:requires": [ "res:ce-84958217-1323-46fe-87c8-652120800190",
"res:ce-8aa6b5b5-095c-4152-89b2-06438129002c" ]
       },{
          "id": "res:ce-84958217-1323-46fe-87c8-652120800190", // List of competencies acquired
"type": "ceasn:Competency",
"ceasn:codedNotation": "EXKSCI1",
"ceasn:competencyLabel": "Knowledge of Science",
"ceasn:competencyText": "Demonstrate a good knowledge of science…"
},{
"id": "res:ce-8aa6b5b5-095c-4152-89b2-06438129002c",
"type": "ceasn:Competency",
"ceasn:codedNotation": "EXKART1",
"ceasn:competencyLabel": "Knowledge of Arts",
"ceasn:competencyText": "Demonstrate a good knowledge of arts…"
```

Dans JSON-LD
https://tinyurl.com/2twaddyv

Principaux points à retenir

- Le CTDL peut être utilisé avec d'autres langages en RDF pour fournir des données détaillées sur les résultats scolaires et professionnels des individus, soit

un graphe de connaissances personnelles

- qui peut être compatible avec des données provenant d'autres sources, telles que des offres de formation, des offres d'emploi, etc.

Pour plus de détails, consultez notre document de discussion :
Les termes CTDL et les données du Registre dans les Verifiable Credentials (VC / Open Badges 3.0 / CLR 1.1)

Merci

Phil Barker
Consultant en systèmes d'information pour l'éducation et la technologie pour améliorer l'apprentissage avec Cetis LLP.
Twitter : @philbarker
E-mail : phil.barker@pjjk.co.uk
Web : http://people.pjjk.net/phil

Site Web : credentialengine.org
Informations techniques : credreg.net
Credential Finder : credentialfinder.org

Retrouvez @CredEngine sur les réseaux sociaux

Ce travail est sous licence Attribution 4.0 International de Creative Commons.