Семантична павутина, поєднані дані, діаграми знань

Філ Баркер, Cetis LLP and Credential Engine
Біг дата для аналізу ринку праці
Он-лайн програма для розвитку можливостей

16 листопада 2022 року

We encourage reuse and remix of this resource with attribution

Credential Engine™

www.credentialengine.org

Теми

- Різноманітність вітається.
- Сумісність – це необхідність.
- Вибір техніки має наслідки.

Це великі теми, які повинні розглядатися у широкому розумінні, але деталі мають значення.

Контекст: Різноманітність та вимога сумісності

Картина ринку праці

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

Різноманітність

Учасник процесу	Інтереси	Взаємодіє з
Роботодавці	Заповнення робочих місць потрібними людьми, найм, просування по службі, нарахування заробітної плати…	Працівники, кандидати на роботу, інші роботодавці, професійні організації, уряд…
Навчальні заклади	Складання програм, організація навчальних курсів, оцінка, акредитація…	Студенти, кандидати на роботу, уряд, організації з забезпечення якості …
Професійні організації	Кваліфікаційні рамки, розвиток кар’єри …	Індивіди, роботодавці, навчальні заклади, організації з забезпечення якості …
Організації з забезпечення якості	Акредитація розроблених навчальних програм та оцінка компетентностей	Навчальні заклади та професійні організації
Уряд	Економіка, законодавство, регламенти…	Всі

Ключові моменти

- Ринок праці включає різних його учасників, кожен з яких має свої власні зацікавленості, практику та способи використання даних.
- Забезпечення сумісності та взаємодії буде полегшено, коли використовуються спільні стандарти.
- Спільні стандарти є можливими, коли є спільні зацікавленості, практики та способи використання даних.

Немає (і не повинно бути) спільного стандарту для всього ринку праці

Однак, є потреба в обміні даними.

Технічна основа семантичної павутини, поєднані дані та діаграми знань

RDF: Рамки опису ресурсу

- Рекомендації W3C (Консорціум всесвітньої павутини) від 1999 року.
- На початку 200-х фокус на «**семантичній мережі**»: онтологія, презентація знань та логічний висновок.
- З 2006 року фокус на «**поєднаних даних**» та сумісності.
- З 2012 року фокус на «**малюнки знань**» та аналізі даних.

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

RDF: Рамки опису ресурсу

Підмет – присудок – Додаток

Все може бути ідентифіковане за допомогою Уніфікованих ідентифікаторів ресурсів :

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

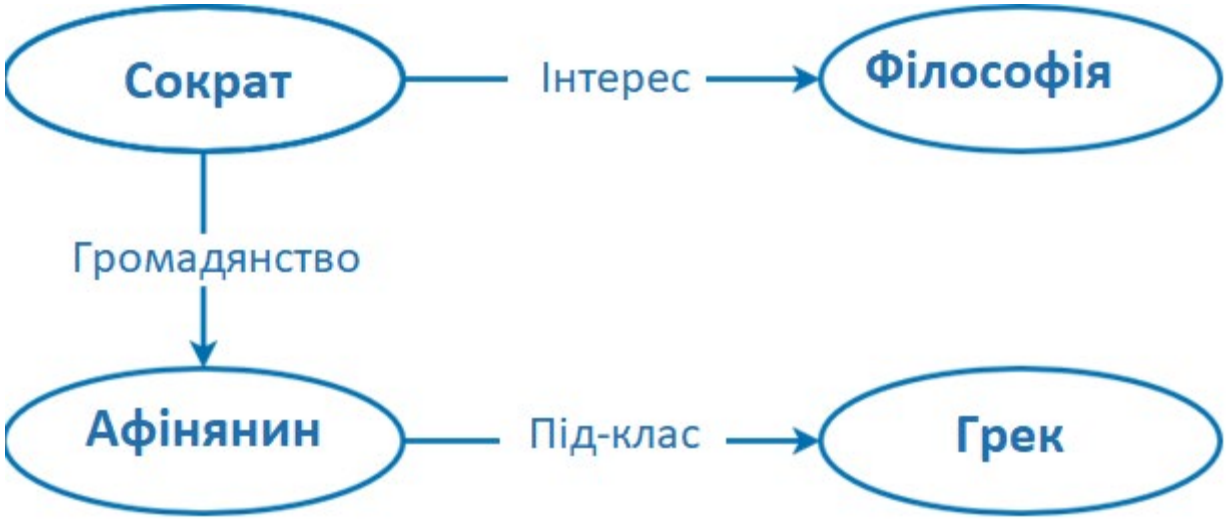
«Люди сприймають Рамки опису ресурсів (RDF), як щось болюче, оскільки вони складні. Але правда – ще гірша. Рамки опису ресурсів болісно спрощені, але вони даютьВам змогу працюватиз реальними даними та проблемами, що єстрашенно складно."

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

Рамки опису ресурсів та семантика

Отже: Сократ - грек

Графіки знань



- Додайте більше даних
- Додаток в одному реченні може бути підметом в іншому.
- З’єднайте речення у вигляді малюнка.
- Почніть давати відповіді на питання типу «скільки греків цікавляться філософією»

Уніфікована ідентифікація ресурсів та поєднані дані

- Визначайте все лише за допомогою Уніфікованої ідентифікації ресурсів

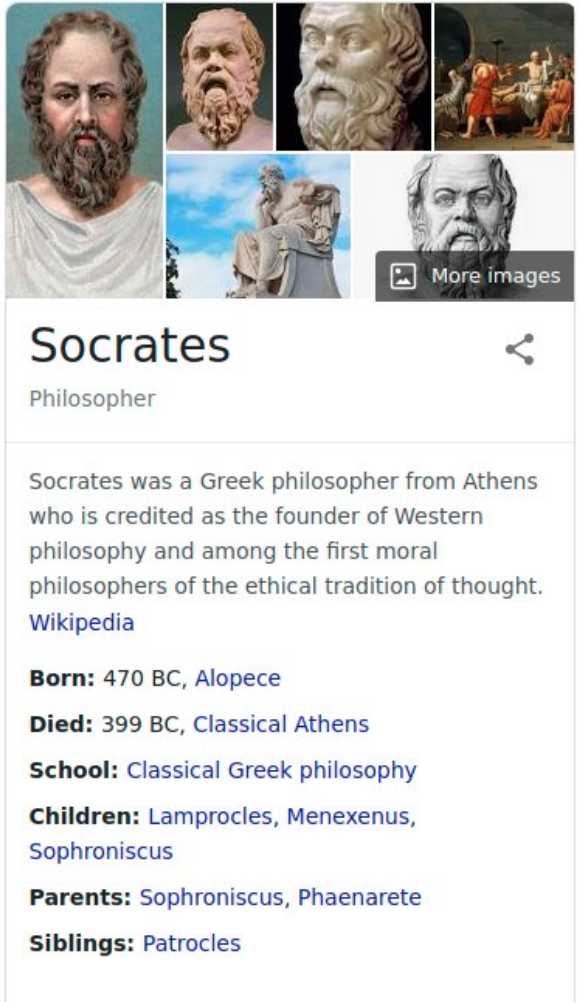


Image is the Google "knowledge panel" from a search for Socrates.

Уніфікована ідентифікація ресурсів та поєднані дані

- Визначайте все лише за допомогою Уніфікованої ідентифікації ресурсів
- Дозволяє уникнути неоднозначності
  - Q: коли Сократ не грек?
  - A: коли Сократ є Sócrates <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

Image is the Google "knowledge panel" from a search for Socrates.

Уніфікована ідентифікація ресурсів та поєднані дані

- Визначайте все лише за допомогою системи Уніфікованої ідентифікації ресурсів
- Дає можливість уникнути неоднозначності
  - Q: коли Сократ не грек?
  - A: коли Сократ – це Со́крат
- Дає можливість поєднання з більшою кількістю даних
- Людська мова незалежна

Image is the Wikidata entry for Socrates name in various languages.

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

Підсумок: Стандарти та сумісність

Якщо організації використовують схожі дані для схожих речей і вони хочуть співпрацювати, вони можуть використовувати однакові стандарти даних для обміну даними (взаємодії)

- Відділ кадрів може надсилати інформацію про вакансії партнеру, що займається працевлаштуванням
- Видавець може надавати дані про навчальні ресурси школам
  В кожному сфері системи IT застосовують відповідні стандарти.
  АЛЕ: різні стандарти підходять для різних сфер.

- Стандарти відділу кадрів для опису заробітної плати та знань необхідних для роботи
- Стандарт системи управління навчанням для опису навчальних курсів та отриманих знань

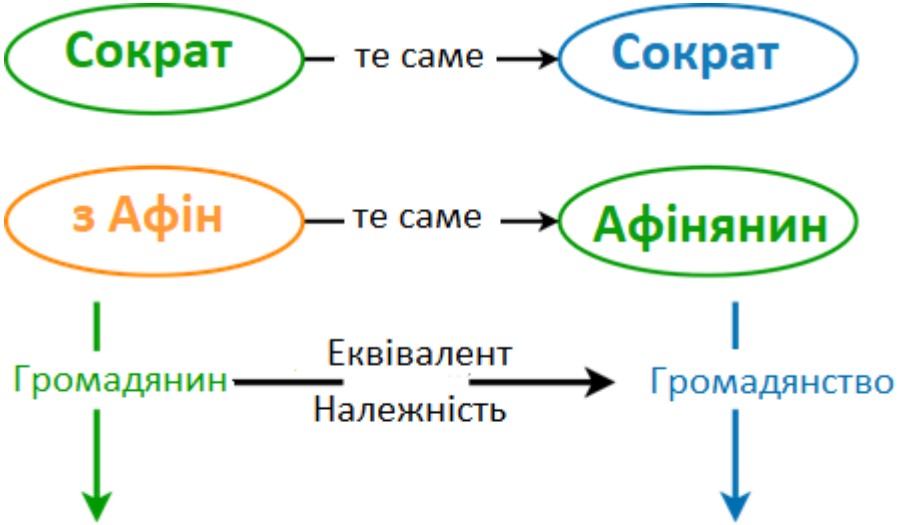
Рамки опису ресурсів (RDF) та гармонізація

Різні стандарти Рамок опису ресурсів (RDF) використовують однакову **мета-модель**

(тобто, повідомлення, як конструкція з трьох складових: підмет – присудок – додаток, визначена за допомогою системи Уніфікованої ідентифікації ресурсів (URLs.))

Тут зелене, помаранчеве та блакитне повідомлення представляють дані з трьох різних стандартів Рамок опису ресурсів (RDF).

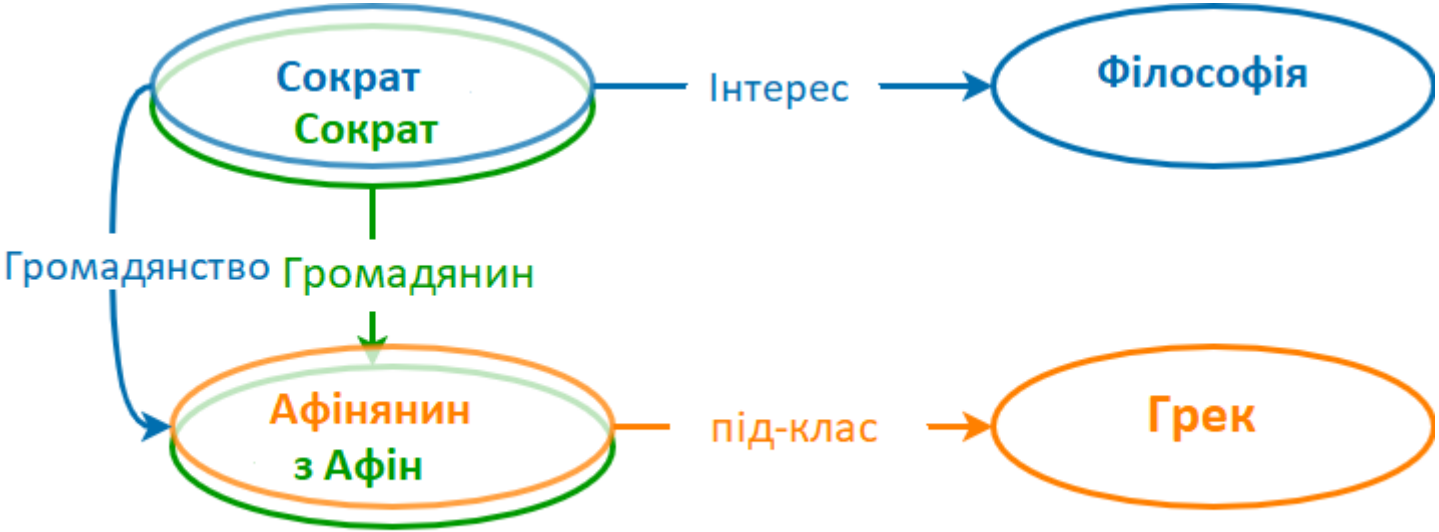
Рамки опису ресурсів (RDF) дозволяють поєднувати дані та знаходити відповідності одних до інших



Терміни з різних стандартів можуть використовуватися разом.

Рамки опису ресурсів (RDF) мають стандарти, щоб установлювати еквівалентність між термінами.

Рамки опису ресурсів (RDF) дозволяють поєднувати дані та знаходити відповідності одних до інших



- Такий само малюнок знання, як і раніше, але створений за допомогою поєднання та відповідності термінів з різних стандартів Рамок опису ресурсів (RDF).

Ключові моменти

Рамки опису ресурсів (RDF):

- Стабільні, зрілі, відкриті рамки стандартів
- Створення семантичної павутини/ поєднаних даних/ малюнків знань
- Дають можливість визначити інформацію, яка не чітко зазначена в даних
- Децентралізовані: зв’язок між наборами даних
- Можливі широкі, відкриті набори даних
- Гармонізація на спільній мета-моделі: сумісність між різними стандартами доменів

Механізм облікових даних:
Прозорість облікових даних, Мова опису прозорості облікових даних (CTDL) та Реєстр облікових даних

Механізм облікових даних та що він робить

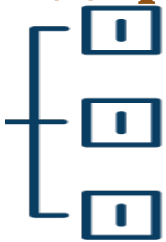
Наша місія: Механізм облікових даних – це неприбуткова місія, яка полягає у створенні карти облікових даних з зрозумілою та повною інформацією, сприянні створенню ресурсів, що дають можливість людям знаходити найкращий для них шлях.

Наше бачення: Ми бачимо майбутнє, де мільйони людей по всьому світу мають доступ до інформації щодо облікових даних, яка відкриває їм очі на широкий спектр можливостей для навчання, розвитку та осмисленої кар'єри.

Як ми реалізуємо нашу місію та бачення:

- Побудова спільної інфраструктури та мови
- Співпраця
- Надання можливостей

Відкриті технології Механізму облікових даних



Мова опису прозорості облікових даних (CTDL)

Схема поєднаних відкритих даних, що описує понад 800 характеристик облікових даних, знань, професій та провайдерів.

Засоби публікації CTDL

Провайдери застосовують Інтерфейс прикладного програмування (API) або інші засоби публікації, що перетворюють інформацію у CTDL, публікують у Реєстрі та створюють Загально єдині ідентифікатори (CTIDs).

Реєстр облікових даних

Реєстр збирає та поєднує облікові дані, дані про знання та професії описані за допомогою CTDL у вигляді графіків поєднаних відкритих даних на допомогу засобам та додаткам, включаючи Пошук облікових даних.

CREDENTIAL FINDER

Пошук облікових даних

Інструмент для перегляду та використання інформації, що зберігається у Реєстрі.

Technical site: https://credreg.net

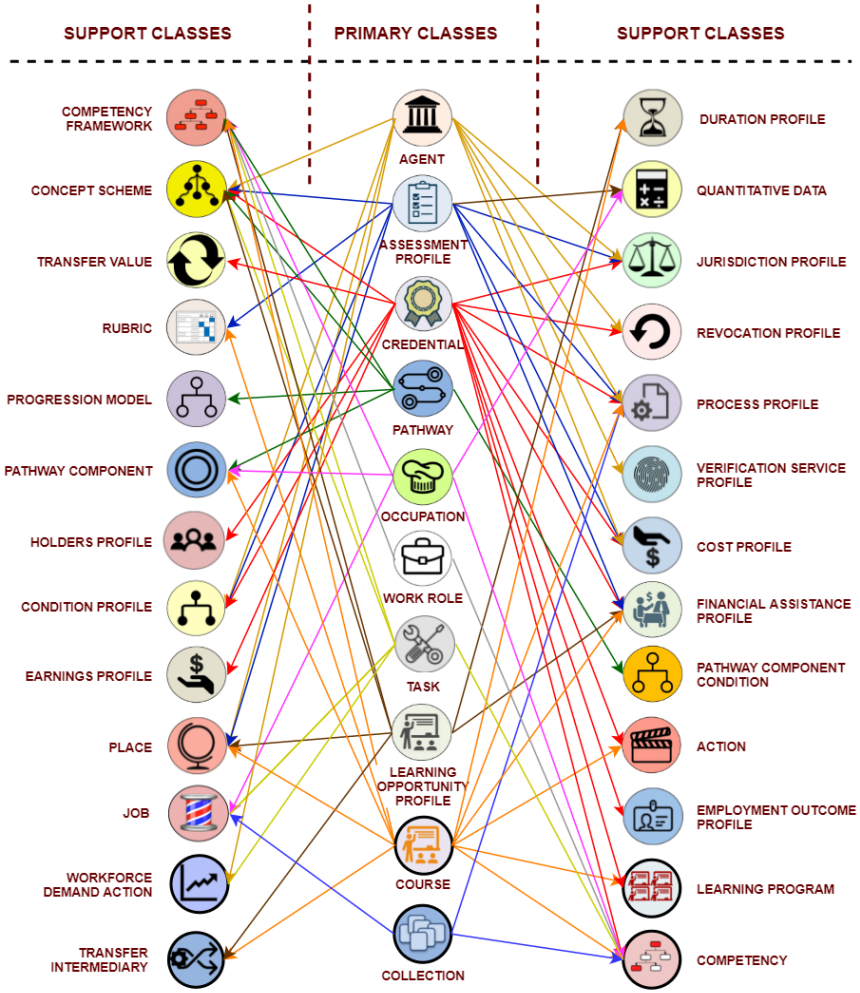
Мова опису прозорості облікових даних (CTDL)

Відкрито ліцензована спільна мова, змодельована на основі специфікацій Консорціумі всесвітньої павутини (W3C) для семантичної павутини, щоб забезпечити прозорість та порівнювальність облікових даних та компетентну інформацію.

Три схеми:

1. CTDL – Облікові дані та суміжні класи
2. CTDL-ASN - Компетентності
3. QDATA – Кількісні дані

CTDL Handbook via https://credreg.net



Терміни Мови опису прозорості облікових даних (CTDL)

Понад 800 термінів поєднаних відкритих даних, що дають багатий опис облікових даних та компетентностей

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

Сертифікація:
Свідоцтво обмежене в часі, яке може бути відкликаним та поновленим органом відповідальним за демонстрацію знань, навичок, та можливостей для виконання конкретних завдань або професійних обов'язків

Приклад з системи Пошуку облікових даних

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

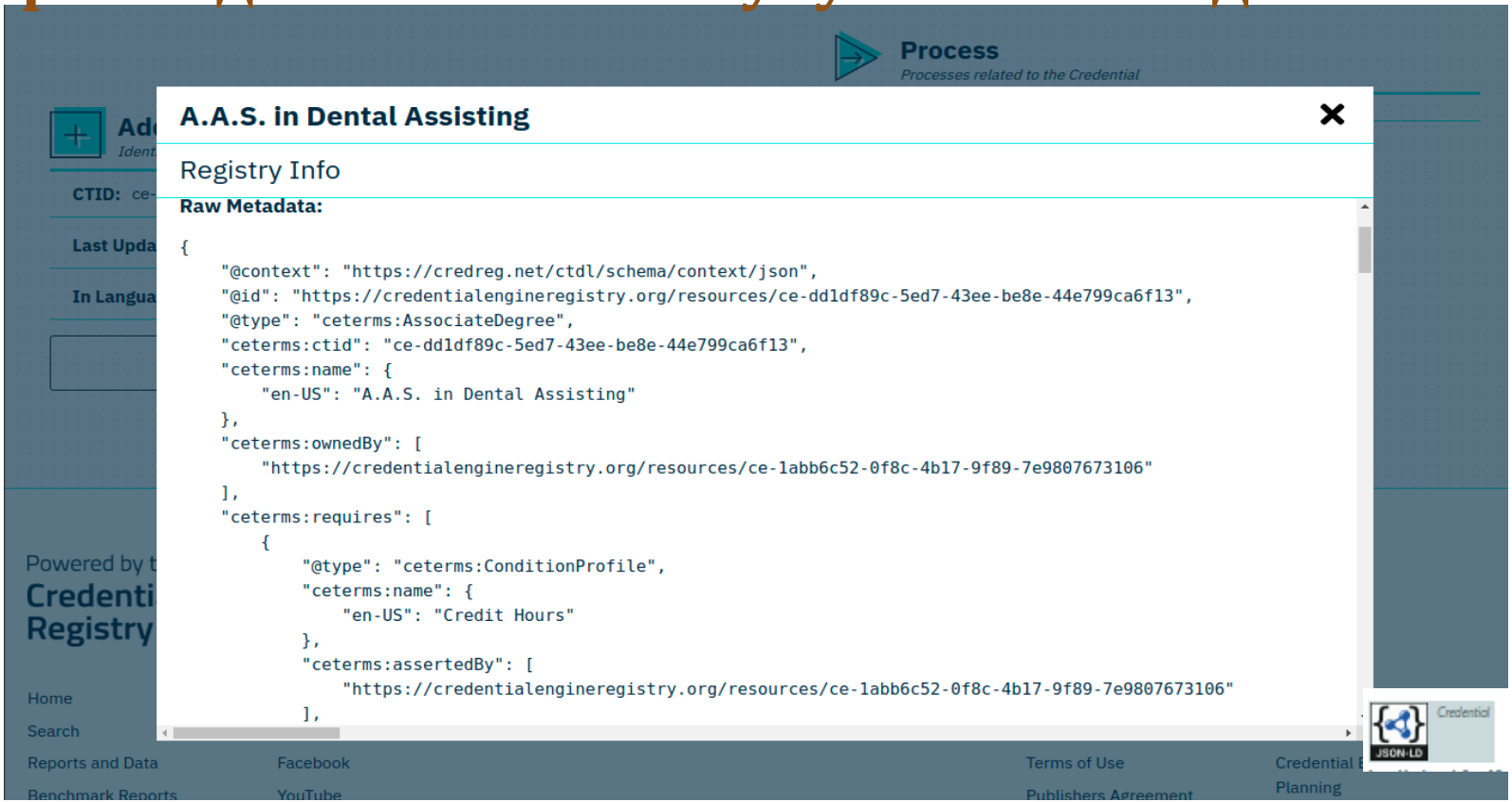
Relationships between the Credential and other data in the Registry

Related Competencies

Teaches 50 Competencies

Вся інформація про облікові дані взята прямо з credentialfinder.org

Приклад з системи Пошуку облікових даних



Вся інформація про облікові дані взята прямо з Реєстру облікових даних, як JSON-LD

Ключові висновки

- Механізм облікових даних є неприбутковою організацією націленою на співпрацю для покращення перспектив освіти та працевлаштування
- Мова опису прозорості облікових даних (CTDL) є мовою, що базується на Рамках опису ресурсів для опису облікових даних (кваліфікацій) та речей пов’язаних з ними.
- Дані, що використовуються Мовою опису прозорості облікових даних (CTDL) можуть зберігатися в реєстрах таких, як Реєстр облікових даних

Мова опису прозорості облікових даних (CTDL) та перевірені облікові дані (коротко)

Консорціум всесвітньої павутини (W3C) Перевірені облікові дані

- Стандарт для очевидних підробок інформації про досягнення особи.
  - Модель даних Рамок опису ресурсів (RDF) https://www.w3.org/TR/vc-data-model/
  - Протокол щодо дотримання приватності інформації, що міститься у цих даних.
- Навчальні приклади в стандарті дуже скупі на деталі.
- Група спільноти Консорціумі всесвітньої павутини (W3C) розглядає випадки використання в освіті.
- На основі 1EdTech(*) OpenBadges V3

(* was IMS)

Приклад перевірених облікових даних Консорціуму всесвітньої павутини (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"
    }
  }
}
```

Недостатньо просто сказати, що у Вас є ступінь.

Мова опису прозорості облікових даних (CTDL) у перевірених облікових даних Консорціуму всесвітньої павутини (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…"
```

In JSON-LD playground
https://tinyurl.com/2twaddyv

Ключові висновки

- Мова опису прозорості облікових даних (CTDL) може використовуватися для інших мов Рамок опису ресурсів (RDF), що надасть повні дані про навчальні та професійні досягнення особи, **графік персональних знань**
- Що може бути сумісним з даними з інших ресурсів таких, як пропозиції навчання, оголошення про роботу тощо.

Для отримання додаткової інформації ознайомтесь з нашим документом про обговорення:
CTDL terms and Registry data in Verifiable Credentials (VC / Open Badges 3.0 / CLR 1.1)

Дякую!

Credential Engine™

Філ баркер
Консультант з інформаційних систем для навчання та технологій для покращення навчання з Cetis LLP.
twitter: @philbarker
email: phil.barker@pjjk.co.uk
web: http://people.pjjk.net/phil

cetis
a co-operative consultancy

Website: credentialengine.org
Technical Information: credreg.net
Credential Finder: credentialfinder.org

Find @CredEngine on Social Media

This work is licensed under the Creative Commons Attribution 4.0 International License.