Data Integration for the Individual using the Solid Platform

May 21, 2019

# Introduction

The Solid platform aims to create a web experience consistent with the original idea that “not only could everyone read content on the web, but they could also help create it. It was to be a collaborative space for all mankind [1].” To create this experience, Solid has addressed the issues of “read-write capability and managed data accessibility” using WebID [2] and Linked Data [3] capabilities. WebID uniquely identifies an individual, while Linked Data is the underlying best practice of “connecting pieces of data, information, and knowledge on the Semantic Web” using a distributed network of URLs. Thus, data can exist anywhere on the web as long as it is accessible via a URL.

The Solid platform is ontology-agnostic but has utilized common ontologies such as Friend of a Friend (foaf) [4], DBpedia [5], and the Web Annotation Ontology [6] for representing data. To further enhance the capabilities of the Solid platform, the ability to connect information horizontally across domains and vertically across levels of specificity using a common ontology would open up additional collaboration and sharing opportunities between individuals and organizations. A well-structured ontology based on standards and best practices would provide not only a list of classes and properties to represent information but would also provide a baseline from which users can extend the ontology to new domains. The Common Core Ontologies [7] (CCO) which is based on Basic Formal Ontology [8] (BFO) provide a rich set of mid and upper ontologies that can fulfill this role of cross-domain information representation that will enable the ability of individuals to share and leverage information relevant to them. It is infinitely extendable and supports querying information at multiple levels so that the degree of exactness vs. generality required by a given use case can be applied.

Our team has investigated the benefits of using CCO with Solid. The team has prototyped a set of capabilities and a simple use case to show that CCO is compatible with Linked Data and offers the potential for enhanced data integration for Solid. The CCO can represent classes and properties that are currently represented using ontologies such as foaf [4], as well as information in other domains of interest. The team is working on documentation that will provide information on how to map the other ontologies of Solid into CCO and will assist interested parties in doing so.

The components of this prototype are described below with a link to where they can be accessed:

1. Common Core Ontologies (<https://github.com/CommonCoreOntology/>): The CCO are open-source, can be extended, and are freely available for adoption by any user. These are sufficiently expressive and extendable to cover any domain. The CCO form the basis of the ontology described below.
2. Common User Profile (CUP) (<https://github.com/CommonCoreOntology/CommonUserProfile>): This is an ontology extended from CCO that contains information describing an individual. The CUP in this prototype is based on a “John Doe” and contains a limited amount of information for the purposes of supporting a data integration use case. All data used in this prototype is stored in a public Solid Pod, but in future prototypes the team will also store private information to demonstrate data security. The CUP will be extended ad infinitum by developers as they utilize and add to its capabilities.
3. CCO Server (<http://www.ontologyrepository.com>): This server resolves CCO classes and properties enabling Linked Data applications to access the raw schema files.
4. Example Alarm Application (<https://github.com/avatar-use-cases/CCO-LD-Avatar>): To show an example of data integration using Solid and CCO, the team has built an application that demonstrates how multiple pieces of data represented as Linked Data triples conformant to the CUP are easily integrated. In this case, the application calculates when a person needs to leave their residence to arrive at a destination by a desired time. It takes into account the amount of time for the individual to get ready, traffic, weather, and commuting preference.
5. CCO-Linked Data Library (<https://github.com/avatar-use-cases/CCO-LD-Avatar>): To enable new data sets to be transformed into CCO-Linked Data triples, a JavaScript library has been written that post-processes Karma [9] mapped data, and loads it into an RdfLib [10] store and Solid Pod. This data is then available to any Solid application.
6. CCO API (<https://github.com/avatar-use-cases/CCO-LD-Avatar>): To assist the developer of Solid applications with the process of developing SPARQL [11] queries for new applications and data integration solutions, an API has been written that ingests CCO-Linked Data triples, interrogates them to determine the links between data elements, and develops a SPARQL query for the user based on the user’s selected data projection and restrictions.

It is recommended that the Solid community will look closely at the benefits of this data integration solution and consider trying it out. We will provide support (especially with the ontologies) for anyone giving this a try.

# Contacts

Please contact the individuals below with questions or requests for support:

Jacob McConomy (US Army CCDC C5ISR), Solid forum ID: JacobMcConomy

Robert Ganger (CUBRC, Inc.), Solid forum ID: rganger

James Schoening (US Army CCDC C5ISR), Solid forum ID: JimSchoening

# References:

[1] “Home | Solid.” [Online]. Available: https://solid.inrupt.com/. [Accessed: 21-May-2019].

[2] “WebID - W3C Wiki.” [Online]. Available: https://www.w3.org/wiki/WebID. [Accessed: 21-May-2019].

[3] “Linked Data | Linked Data - Connect Distributed Data across the Web.” [Online]. Available: http://linkeddata.org/. [Accessed: 21-May-2019].

[4] “FOAF Vocabulary Specification.” [Online]. Available: http://xmlns.com/foaf/spec/. [Accessed: 21-May-2019].

[5] “Ontology | DBpedia.” [Online]. Available: https://wiki.dbpedia.org/services-resources/ontology. [Accessed: 21-May-2019].

[6] “Web Annotation Vocabulary.” [Online]. Available: https://www.w3.org/ns/oa. [Accessed: 21-May-2019].

[7] “CommonCoreOntology,” *GitHub*. [Online]. Available: https://github.com/CommonCoreOntology. [Accessed: 21-May-2019].

[8] “Basic Formal Ontology (BFO) | Home.” [Online]. Available: https://basic-formal-ontology.org/. [Accessed: 21-May-2019].

[9] *Information Integration Tool. Contribute to usc-isi-i2/Web-Karma development by creating an account on GitHub*. USC ISI I2, 2019.

[10] *Linked Data API for JavaScript. Contribute to linkeddata/rdflib.js development by creating an account on GitHub*. Read-Write Linked Data, 2019.

[11] “SPARQL,” *Wikipedia*. 28-Feb-2019.