Enable Decentralised Semantic Content Negotiation through Equivalence Links

Yousouf Taghzouti

Enable Decentralised Semantic Content Negotiation through Equivalence Links

Yousouf Taghzouti[0000−0003−4509−9537]

Mines Saint-Etienne, Univ Clermont Auvergne, INP Clermont Auvergne, CNRS, UMR 6158 LIMOS, F - 42023 Saint-Etienne France
yousouf.taghzouti@emse.fr

Introduction

The Web is a decentralised system where servers can serve a set of URIs that identify resources. A resource may have one or more representations [2]. This encourages Content Negotiation (CN), the mechanism by which a client can request a resource representation that satisfies some constraints [1, Section 12]. In the Semantic Web, resources are described using different vocabularies. HTTP provides the means to negotiate representations that have a required media type. Similarly, semantic validation languages (e.g. SHACL) could be used to define the constraints that knowledge graphs must satisfy. If a resource were identified by a single URI (*unique names assumption*), this would imply that all representations would be on one server. However, this is not possible on the current Web, because Web standards do not assume unique names and representations are distributed in different places. We propose an approach to perform content negotiation even when representations are distributed and present in multiple locations by using *equivalence links*, which involves on-the-fly SHACL shape validation.

Decentralised Semantic Content Negotiation Problem

Content Negotiation is the mechanism that enables the selection of a representation from among multiple ones available under the same URI. *The unique names assumption* states that: different names in the world refer to different things. On the Web, this would imply that different URIs should be used to identify different resources. Therefore, a resource identified by a URI would have all its possible representations on a single server. As a result, a client would be able to negotiate a desirable representation of a resource with that server. However, on the actual Web, representations are distributed across different servers, and a URI can only be served by a single one. Consequently, when negotiating with a server, only a subset of all existing representations is considered. Similarly, the Web of Linked Data (the largest publicly available Knowledge Graph) is inherently distributed, and the unique names assumption does not hold. Fortunately, equivalent links such as *owl:sameAs* can be used to indicate that two URIs actually refer to the same entity, but they have not yet been used in the CN flow.

Employ CN in a Decentralised Semantic Context

In the following, we will explore our approach in two negotiation dimensions: *media type* and *profile* [6, Section 3.1]. In the first case, the server expects a client to request a resource identified by a URI and to provide a set of constraints (in this case, a set of acceptable media types). The client expects to have a representation (web document) that validates the constraints, and possibly a set of plausible alternative URIs to continue the negotiation if desired. The negotiation starts with the server trying to check if any of the representations are plausible. If not, it first looks for a set of sameAs URIs. It then either (1) redirects the client to these other potential representations by sending a *300 (MULTIPLE CHOICE)* [1, Section 15.4.1] response, or (2) proceeds one step further to check whether any of the alternatives discovered are valid. If no representation could be found, only then will the server respond with an error *406 (NOT ACCEPTABLE)* [1, Section 15.5.7]. The same idea could be carried over from the Web of documents to the Web of data, but with some modifications to adapt it to the context of the Semantic Web and the negotiation in the profile dimension [5]. First, the server assumes that the client is negotiating RDF documents, and therefore constraints must be defined for this type of document (e.g. using SHACL [3] or ShEx [4]). The server keeps track of the best graph and uses the *isABetterRepresentationThan* function to test whether an alternative valid graph is better than the current best graph. The function takes two graphs and produces a Boolean evaluation, e.g. the evaluation could be based on the number of nodes, a graph with more nodes potentially holds more knowledge.
Implementation We have developed a Java implementation of our approach. Our prototype[1] is built using the *Spring framework*, we provide two endpoints `</dcn/api/media-type>` and `</dcn/api/profile>`. The resource URI is provided using the *iri* query parameter, and the constraints are passed in a header-based manner, *accept* and *accept-profile* for the media type and profile dimensions respectively. The *sameAs service*[2] is used to obtain the set of equivalent URIs. *SHACL* (the W3C recommendation) is used to express the constraints that RDF sources must conform to, while *Apache Jena* is used to manipulate the RDF graphs and perform validation. *Swagger* is used to provide friendly API documentation. Figure 1 shows an example of a request.

Conclusion

In this document we present an approach to achieve decentralised content negotiation using Semantic Web technologies. We use equivalence links such as *owl:sameAs* to discover potential representations. The conclusion is that the use of such links, present in knowledge graphs, enables more effective content negotiation of Web resources by allowing the discovery, validation and serving of representations stored in a distributed manner, contributing to an overall increase in the availability of resource representations.

[1] Github repository: https://github.com/YoucTagh/decentralised-cn
[2] SameAs service home page: http://sameas.org/

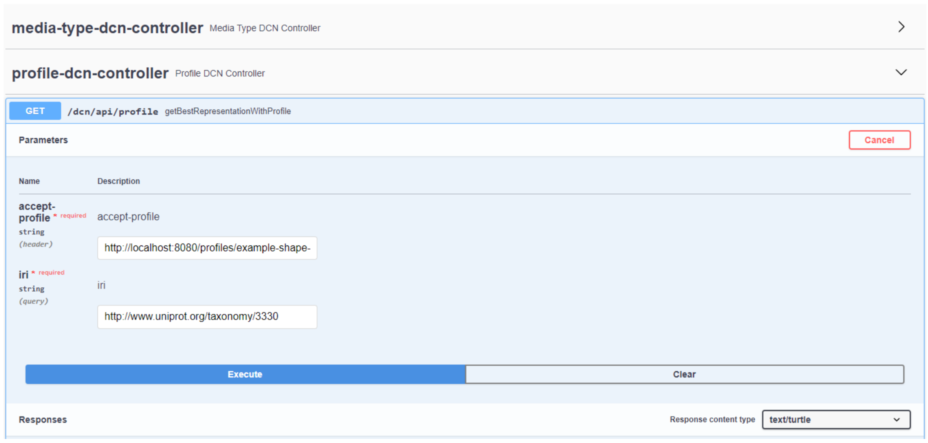

Fig. 1: A request sent using the Swagger-UI to get a representation of a resource that validates a SHACL shape graph.

References

1. Fielding, R.T., Nottingham, M., Reschke, J.F.: HTTP Semantics. RFC 9110, IETF (Jun 2022), http://www.ietf.org/rfc/rfc9110.txt
2. Jacobs, I., Walsh, N.: Architecture of the World Wide Web, Volume One, W3C Recommendation 15 December 2004. W3C Recommendation, W3C (Dec 15 2004), http://www.w3.org/TR/2004/REC-webarch-20041215/
3. Knublauch, H., Kontokostas, D.: Shapes Constraint Language (SHACL). W3C Recommendation, W3C (Jul 20 2017), https://www.w3.org/TR/2017/REC-shacl-20170720/
4. Prud'hommeaux, E., Boneva, I., Labra Gayo, J.E., Kellogg, G.: Shape Expressions Language 2.1. W3C Community Group Report, W3C (2019), http://shex.io/shex-semantics-20191008/
5. Svensson, L.G., Atkinson, R., Car, N.J.: Content Negotiation by Profile. W3C Working Draft, W3C (Nov 26 2019), https://www.w3.org/TR/2019/WD-dx-prof-conneg-20191126/
6. Taghzouti, Y., Zimmermann, A., Lefrançois, M.: Content Negotiation on the Web: State of the Art. Tech. Rep. abs/2204.10097, CoRR (2022), https://arxiv.org/abs/2204.10097