

Domain-Specific Backlinking Services in the Web of Data

Manuel Salvadores*, Gianluca Correndo*, Martin Szomszor[†],

Yang Yang, Nick Gibbins, Ian Millard*, Hugh Glaser*, and Nigel Shadbolt*

* IAM Group, School of Electronics and Computer Science, University of Southampton, UK

[†] City eHealth Research Centre, City University London, UK

{ms8,gc3,yy1402,nmg,icm,hg,nrs}@ecs.soton.ac.uk, martin.szomszor@city.ac.uk

Abstract—This paper describes an Open Linked Data *backlinking* service, a generic architecture component to support the discovery of useful links between items across highly connected data sets. Using Public Sector Information (PSI) currently available as Linked Data, we demonstrate that contemporary publishing practices do not adequately support the ability to navigate or automatically traverse between resources published by different vendors, or the capacity to discover information relevant to a particular URI. Although some useful services in this area have been developed, such as large triple indexes of published data, and the collection of *sameAs* relationships between individuals, we believe that an important component is missing: a mechanism to discover the backlinks to relevant resources that cannot be found by direct URI resolution. We present the implementation of such a component, integrating data from various PSI sources.

I. INTRODUCTION

The Open Linked Data (LOD) Initiative represents a collaborative effort to publish and link large amounts of data over the web using standard protocols and agreed representations. Much of this effort is centered on the premise that URIs are resolvable [2] and return RDF data that subsequently links it to other data items, creating a large *Web of Data* (WoD). In the formative years, Linked Data publishing was rolled out across a variety of domains, including entertainment, science, encyclopedia, government data, etc. It is anticipated that once the number of items exposed reaches a critical mass, reuse of data will become increasingly popular, fostering a new paradigm for the publishing and consumption of data, much like the Web 2.0 revolution did for REST [7], XML, and SOAP. In anticipation of this moment, our attention is drawn towards the additional architectural components that must be in place to support the consumers of Linked Data.

In a recent LOD exercise [11], we built an application to integrate data from various public sector information sources, such as the UK Government, the UK Ordnance Survey¹ and DBPedia². The aim was to present an overview of a geographical location in terms of the crime, health, education, and traffic statistics associated with it. During this exercise, we found that following only LOD principles

of i) resolve URI, ii) parse data, iii) follow links to other URIs, often meant we missed crucial joining data that would allow us to integrate information from different sources.

Although in some cases it is possible to discover additional locations in which a URI is used, for example by using the void Vocabulary [1] to discover SPARQL endpoints that can be queried, in the general case it is not possible using URI resolution alone because links are only uni-directional. If an RDF graph resolved from URI_1 contains a reference to URI_2 , but the graph from URI_2 has no reference to URI_1 , it is impossible to discover the relationship between the two URIs by resolution if only URI_2 is known.

In this paper, we present our Backlinking Service, a generic component to support the publishing and discovery of useful links between URIs from different Linked Data publishers using only URI resolution. We begin in Section II with a review of the current state-of-the-art. Section III focuses on the backlinking problem with respect to Linked Data and presents a motivating example for why a backlinking service is required. Section IV describes our implementation of a backlinking service and how it can be integrated with a co-reference service to maximize its utility. Our contribution is evaluated in Section V where the benefit of our backlinking service is tested with UK PSI data. In Section VI our conclusions and future work are presented.

II. BACKGROUND

The World Wide Web (WWW) is a directed graph of documents (i.e. links that connect nodes in the hypergraph are unidirectional) and as such is asymmetric. This influences both its navigability and the possibility of information discovery. Historically some of the first hypertext systems implemented bidirectional links to freely enable navigation in both directions (e.g. Enquire, Microcosm [8] and Xanadu [10]). The possibility of having backlinks was also taken into account in the design of link topology³ in the infancy of the WWW and since then proposals have been made for building such links off line [4].

There are already a few proposals born within the hypertext community for gathering backlinks (e.g. [3] or the

¹<http://www.ordnancesurvey.co.uk> The mapping agency for Great Britain

²<http://dbpedia.org> a Semantic Web representation of Wikipedia

³<http://www.w3.org/DesignIssues/BuildingBackLinks.html> last access 24/03/2010

Google search engine using the “link” option), but these proposals are not aimed at the Semantic Web, and target the traditional Web of documents. Distributed approaches proposed in the past [13] relied on logging the *Referer* HTTP header (see [6], section 14.36) during the usual document serving activity by a web server. Such information can then be stored in distributed databases and used afterwards.

The Linked Data community is pursuing the use of the Web for publishing and connecting data adopting RDF as a common data model. The main topology of the Semantic Web information network is very similar to the WWW, containing links that connect pieces of information that are discoverable and browsable in only one direction. New technologies have been developed by the Semantic Web community to provide more diverse access to Linked Data sets. SPARQL endpoints are sometimes provided for running queries against data sets, and indexing services for linked open data (e.g. *sindice.com*) can support the discovery of backlinks. *Sindice* [12] is a lookup index over RDF documents crawled on the Semantic Web, and it allows users and software agents to discover documents containing information about a given resource. Such centralized services can be used in order to discover links pointing to a resource by just collecting all pertinent documents, parsing them and recording the triples that explicitly mention the resource of interest. The performance of discovering backlinks in this fashion greatly depends on the size of the documents returned by the lookup process. The greater the number of triples contained in a single document, the heavier the process of parsing and querying it is and this must be taken into account in order to design a scalable service.

Significant help to discover backlinks could be provided by the extensive use of *void*⁴[1] descriptors. *VoiD* descriptors give a precise notion of the topology of the linked data cloud and could be used in order to narrow the rose of possible data sets that could provide incoming links to a resource of interest. However such knowledge should be coupled by the capability, provided by those data sets, to return the incoming links. URI resolution is out of the question because the URIs to resolve cannot be known in advance; they are the very object of this research. Therefore a SPARQL endpoint or a lighter service that can return such links must be implemented by each data set, and “*void* + SPARQL” can be considered a backlinking service. However, we cannot expect all the data sets in the WoD to provide a SPARQL end-point (expensive to maintain) and use *void*. (from the data consumer perspective, not trivial to use).

Semantic Web approaches for discovering backlinks in the Web of Data are still missing. Although there are some tools that can provide backlinks (such as *sindice.com*), these operate at the document level and they do not facilitate the

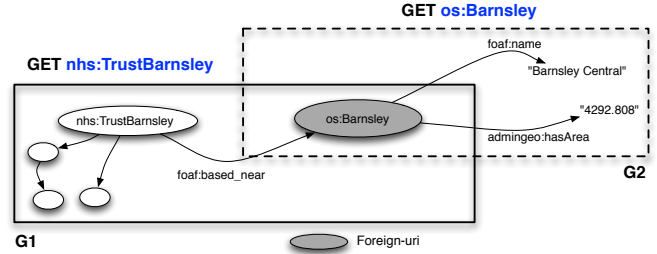


Figure 1. foreign URI (with respect to nhs.psi.enacting.org dataset)

discovery of backlinks at the graph level. This means that current solutions are not adequate for the WoD because their results give back documents but not resolvable URIs.

In contrast to the methods outlined above, in our work we have focussed on the basic tenets of Linked Data in trying to provide a lightweight solution to the problem of retrieving backlinks from a resource of interest, relying on resolvable URIs only. Another import characteristic of our solution is that our backlinking service is itself Linked Data, therefore it is seamlessly integrated and can be seen as a Linked Data layer on top of the current WoD.

III. MOTIVATION

The four Linked Data principles [2] promote RDF as the standard data model and encourages the community to create a decentralized RDF graph at a Web scale, the WoD. We have identified a navigability issue in the WoD and this paper proposes a solution to improve the connectivity of the RDF graph that represents the WoD. The solution we propose is a lightweight service designed to facilitate backward navigation between different data sets. Backward navigations are broken if the object of the triple to be followed comes from a different data set. Figure 1 shows two RDF graphs that can be obtained by resolution of two URIs:

- G_1 by HTTP resolution of `nsh:TrustBarnsley`⁵
- G_2 by HTTP resolution of `os:Barnsley`⁶

`nsh:TrustBarnsley` represents a hospital in the UK National Health Service (NHS), and `os:Barnsley` is a resource that represents a UK geographic location. `os:Barnsley` is a foreign URI in G_1 because its domain is not the same as the publisher (`nhs.psi.enacting.org`). At this point we can say that “`os:Barnsley` is a foreign URI in the graph defined by `nsh:TrustBarnsley`”. In general, we can define the term *foreign URI* as:

“A URI X is considered foreign in a RDF Graph G if exists a triple (s, p, X) in G and $domain(x) \neq domain(G)$ ”

Foreign URIs represent a barrier when navigating the WoD and it is a pattern that happens quite often. This is

⁵http://nhs.psi.enacting.org/id/A_RFF

⁶<http://data.ordnancesurvey.co.uk/id/7000000000024753>

⁴<http://semanticweb.org/wiki/VoiD> last accessed 24/03/2010

due to the fact that there are many statements with *foreign URIs* generated by join point data sets.

A. Broken Navigation with foreign URIs

Foreign URIs make data discovery difficult because it is not possible to navigate the RDF documents of the WoD bidirectionally. Let us assume we want to discover alternative resources or additional facts concerning the URI `dbpedia:Barnsley`⁷. Which happens to be the equivalent `os:Barnsley`. This type of equivalence can be looked up in the WoD through co-reference systems. A co-reference system, such as *sameAs.org*, is a service that enables the discovery of sets of resources that are equivalent. These services increase considerably the connectivity of the WoD facilitating the integration of data sets. *sameAs.org* is a domain-independent service for the WoD. Our work pursues better connectivity in the WoD and we consider that co-reference systems are partially achieving this outcome, and they resolve it partially because co-reference systems do not resolve backward navigation with *Foreign URIs*.

The broken navigation between `nsh:TrustBarnsley` and `dbpedia:Barnsley` gets reflected through the following example. To summarize, the following triples are decentralised in separate graphs of the WoD:

- T_1 (`nsh:TrustBarnsley`, `foaf:based_near`, `os:Barnsley`) Obtained by direct resolution of `nsh:TrustBarnsley`.
- T_2 (`os:Barnsley`, `owl:sameAs`, `dbpedia:Barnsley`) Obtained via RESTful query to *sameAs.org*.
- T_3 (`dbpedia:Barnsley`, `foaf:name`, "Barnsley Central") Obtained by direct resolution of `dbpedia:Barnsley`.

For clarity, let's assume the following symbol equivalences: $X = nsh:TrustBarnsley$, $Y = os:Barnsley$ and $Z = dbpedia:Barnsley$. So if we take T_1 , T_2 and T_3 as a story board to navigate through them, then we realize that we can perform the navigation forward ($X \rightarrow Y \rightarrow Z$) but not backwards ($Z \rightarrow Y \nrightarrow X$). When navigating backwards, starting from Z , X is not reachable by recursive HTTP resolution. This happens because the publisher of Y (`os:Barnsley`) is not meant to return URIs from other data sets that link to its resources. Current practice only recommends to provide backlinks within the same dataset. Therefore we will not find any triples mentioning X in the RDF graph represented by Y , we denote this case by $Y \nrightarrow X$. In this case, the bidirectional navigation can be represented as:

$$(X \rightarrow Y \rightarrow Z) + (Z \rightarrow Y \nrightarrow X) = (X \rightarrow Y \leftrightarrow Z)$$

Even though this backward navigation is broken it is up to the publishers to avoid this issue by keeping all the triples they produce with URIs that belong to their domain. For those cases co-reference systems are an optimal solution to traverse between resources from different data sets.

⁷http://dbpedia.org/page/Barnsley_Central

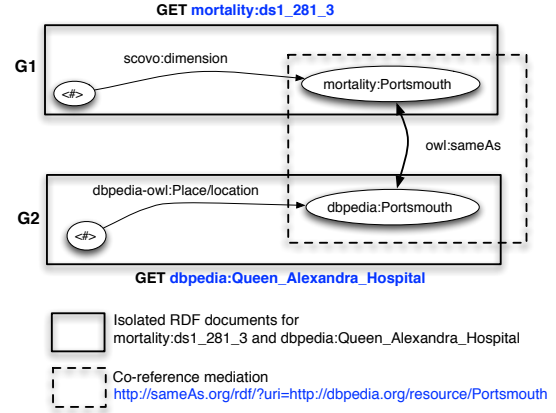


Figure 2. Co-reference Systems Approach

B. Co-reference System Approach

The practice for publishing Linked Data recommends that backlinks are published for the URIs that belong to the same data set⁸. Let us assume two URIs (`mortality:ds1_281_3`⁹, `dbpedia:Queen_AH`¹⁰) produce RDF graphs without *foreign URIs*. For these we will demonstrate that a co-reference service such *sameAs.org* resolves the navigation problem.

`mortality:ds1_281_3` represents mortality statistics in SCOVO[9] for the UK region of Portsmouth, `dbpedia:Queen_AH` is the resource that represents a hospital located in the same geographical region. The RDF graphs retrieved by resolving these URIs are not linked together by themselves because each of the publishers, following the Concise Bounded Description (CBD)¹¹, only returns a subset of the RDF graph within their data sets.

This problem can be solved by using a co-reference system. Both graphs are pointing to URIs that are located in a common bundle in *sameAs.org* therefore the two RDF graphs are now part of a connected RDF graph and it is possible to navigate automatically between them.

Figure 2 shows how the co-reference system connects bidirectionally the isolated RDF graphs. In this case, it is possible to navigate from `mortality:ds1_281_3` to `dbpedia:Queen_AH` and vice-versa by just making recursive HTTP resolutions. This example works based on two assumptions:

- The publishers from both domains describe their resources providing backlinks to the subjects from the same data set following the CBD mechanism.

⁸<http://www4.wiwiw.fu-berlin.de/bizer/pub/LinkedDataTutorial/> last accessed 24/03/2010

⁹http://mortality.psi.enacting.org/id/ds1_281_3

¹⁰http://dbpedia.org/resource/Queen_Alexandra_Hospital

¹¹<http://www.w3.org/Submission/CBD/> last accessed 24/03/2010

- The publishers avoid making statements that include *foreign URIs*.

Taking into account these two assumptions the following example represents a possible story board of HTTP resolutions. Let's assume the following symbol equivalence: $X = \text{dbpedia:Queen_AH}$, $Y = \text{dbpedia:Portsmouth}$, $Z = \text{mortality:Portsmouth}$, $T = \text{mortality:ds1_281_3}$.

Looking at the triples represented in Figure 2, we can see that none of the URIs X , Y , Z and T represented in the graphs G_1 and G_2 are foreign URIs because they all come from the same domain as the graph they belong to. Therefore, following the practice, the resolution of X , Y , Z and T will provide the backlinks within the same data set and using a co-reference service as mediator we can perform bidirectional navigation between any pair of URIs from G_1 and G_2 . We can denote this via:

$$(X \leftrightarrow Y) + (Z \leftrightarrow T) + (Y \equiv Z) = (X \leftrightarrow Y \leftrightarrow T)$$

(or)

$$(X \leftrightarrow Y) + (Z \leftrightarrow T) + (Y \equiv Z) = (X \leftrightarrow Z \leftrightarrow T)$$

Since X and Y are URIs that belong to the same data set we can assume the publisher will return backlinks between them when resolving any of them. The same assumption applies between Z and T . The mediation piece ($Y \equiv Z$) represents the co-reference discovery provided by sameAs.org.

This navigation example relies on the assumption that RDF documents in the WoD do not contain *foreign URIs*. This assumption does not hold in reality because the Linked Data community evangelizes that to realise the power of the WoD one should wherever possible re-use URIs rather than duplicating descriptions of entities or concepts. Therefore we predict that WoD will experience a rapid increase of *foreign URI* linkages as the number of data hubs increase.

The main motivation of this research is to enable the discovery of *foreign URI* linkages by designing and implementing a lightweight *backlinking* service. In order to probe the benefit of such a service we have focussed on a specific domain, the UK Public Sector Information domain.

IV. PUBLIC SECTOR INFORMATION BACKLINKING SERVICE

To facilitate consumers of the WoD in the discovery of *foreign-URIs*, we have designed a backlinking service that keeps track of this special type of link pattern. We have implemented this service for a number of UK PSI data sets currently available as Linked Data. The service, located at <http://backlinks.psi.enacting.org>¹², is designed according to three basic principles:

¹²Refer to <http://backlinks.psi.enacting.org> for usage examples and API documentation

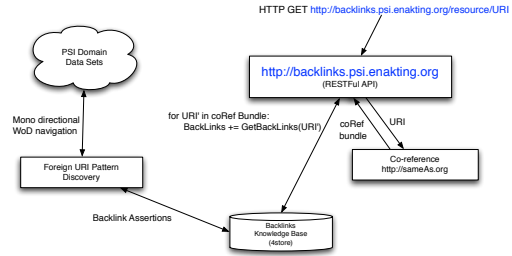


Figure 3. High-level Architecture

- **Lightweight Service**: The service must be easy to use and resolve a specific problem.
- **Linked Data Compatible**: The backlinking service should be compatible with current Linking Open Data best practises, with all requests supporting URI HTTP resolution and content negotiation.
- **Co-reference Support**: to extend the WoD graph coverage the backlinking service is integrated with a co-reference system. This approach allows us to reach data sets from across unbounded WoD.

Figure 3 is a visual overview of the architecture, and is composed of three parts:

- **Foreign URI Pattern Discovery Component**: This is the component responsible for automatically navigating the PSI data sets from the WoD and identifying *foreign URIs*. This component crawls the WoD retrieving all the *foreign URIs* found in the data sets under the study. It resolves all the RDF documents from a starting (or input) list of URIs and inspects each document returned to identify triples in which the object is a *foreign URI*. For every foreign URI found we assert a *rdfs:seeAlso* statement into the knowledge base. The *seeAlso* statement is a triple that points to the original URI in case of backward navigation. For instance, if the service was analysing `nsh:TrustBarnsley` (see Section III) then we would discover that the document returned by resolving that URI contains a triple in which `os:Barnsley` is in the object position, i.e. `os:Barnsley` is a foreign URI in this context. If a client were seeking information about `os:Barnsley` then it may wish to discover the information about that concept contained within the document retrieved when resolving `nsh:TrustBarnsley`. As a result, the corresponding assertion into the backlinks knowledge base is:

```
os:Barnsley
  rdfs:seeAlso nsh:TrustBarnsley .
```

Which follows the pattern:

```
<FOREIGN-URI>
  rdfs:seeAlso <LOCAL-URI>
```

- **RESTful API as Front-end Service**: Access to the

service is provided by a RESTful API that accepts requests by simple HTTP GETs. The interface is: <http://backlinks.psi.enacting.org/resource/URI> Where *URI* is the resource for which we want to discover the backlinks. The service queries the knowledge base where the *seeAlso* statements were asserted and returns a document with all the backlinks. The output of the service can be obtained in JSON, RDF+XML, TURTLE or HTML, either through the Service URL or by specifying the accept header of the HTTP request¹³. The logic of this service is integrated with a co-reference system extending its functionality to all the URIs in a *sameAs* bundle. Let us assume a *sameAs* bundle is made of three URIs:

$$sameAsBundle = \{URI_1, URI_2, URI_3\}$$

and there is a set of two backlink assertions in our system:

$$\begin{aligned} &\{ \langle URI_2 \rangle \textit{seeAlso} \langle URI_x \rangle \} \\ &\{ \langle URI_2 \rangle \textit{seeAlso} \langle URI_y \rangle \} \end{aligned}$$

The system will return $\{URI_x, URI_y\}$ as backlinks when receiving any of the URIs part of the *sameAs-Bundle* as input URI.

- Knowledge Base: We have chosen to use 4store for the internal storage of the backlinks knowledge. While a number of mechanisms could have been employed, using an RDF based store offers significant benefits in terms of flexibility, and we already have an enterprise scale platform available for hosting the service.

The system records backlink metadata, such as the RDFS label and the RDF type(s) of every URI subject of a backlink. Hence, the complete output for a backlink in the system is as follows:

```
os:Barnsley rdfs:seeAlso nsh:TrustBarnsley
nsh:TrustBarnsley a nhs:OrgName;
    rdfs:label "Barnsley Hospital ...".
nhs:OrgName rdfs:label "NHS Organisation".
```

A. Geographic ad-hoc Semantics: an optional capability

Many of the PSI data sets published so far are related to a spatial and temporal dimension, in other words, all data can be linked together by its spatial and temporal indexes. This is unsurprising, the spatial and temporal reasoning have always been considered to be an important part of common-sense reasoning in Artificial Intelligence. Pursuing a better connected WoD we developed a Geographical Service for the WoD. This service computes the closure of partonomies for the UK geography taking as source the Ordnance Survey Linked dataset. To improve the backlinking coverage we aim

¹³More detailed documentation about the API can be found at <http://backlinks.psi.enacting.org>

to get all the possible containments from all the dictionaries supported in the geographical service for a given URI. There is a natural outcome from this integration and it can be shown with the following example: when asking for backlinks connected to the URI `dbpedia:Hampshire`. Prior to the use of the geographical extension a request to retrieve backlinks for `dbpedia:Hampshire` would just give back 14 URIs. This same request when the geographical service is integrated returns 12 345 resources¹⁴ contained within Hampshire.

We have kept the decentralised nature of the Backlinking and Geographical services and the Backlinking Service performs HTTP requests to get the geographic containments. When the geography extension is enabled the Backlinking service gets the list of contained entities for the input URI and returns the backlinks connected to any URI part of those containments. In [5] is described such integration and the implementation of the Geographical Service. This ad-hoc reasoning exploits the semantics of such contextual dimensions for easing entity retrieval and browsing.

V. EVALUATION

Following Tim Berners-Lee's call for *raw data now*¹⁵ the UK Prime Minister decided to favour the opening of Public Sector information. The Public Sector data is retained, up to now, by central government, local councils, the NHS, police and education authorities and other governmental institutions. By releasing and giving open access to this information the government intends to increase its transparency and create economic and social capital. Our evaluation framework is based on data sets that are part of the government initiative, the Ordnance Survey and the EnAKTing project. The backlinking assertions have been sourced from the following data sets: `*.psi.enacting.org`, `data.ordnancesurvey.co.uk` and `*.data.gov.uk`¹⁶

From all the selected data sets we have identified `data.ordnancesurvey.co.uk` as the the hub of link-ages. This means that all the other data sets in some manner (either directly or through *sameAs.org*) link to the UK Geographic data set. All the data sets under the domain `*.psi.enacting.org` use the Ordnance Survey ontology, via the `owl:sameAs` alignments. Other datasets from `*.data.gov.uk` adopt the same ontology as their authoritative source for geographical information.

The presence of authoritative datasets, such as `data.ordnancesurvey.co.uk`, is a strong point for data integration and improves the connectivity of the WoD, but it also weakens the navigation since it generates a

¹⁴Refer to <http://backlinks.psi.enacting.org/help> for a more graphical explanation.

¹⁵http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html last accessed 20/12/2009

¹⁶The extended version of this paper contains statistics of each of this datasets <http://eprints.ecs.soton.ac.uk/21256/>

high level of *foreign URI* statements in RDF documents that reside in other domains. Our backlinking service allows the user or agent to navigate backwards from the authoritative data sets towards these RDF documents with *foreign URIs*.

The provided backlinking service enhances the navigation of the WoD by augmenting the connectivity of the information network. It provides links to resources that were not reachable before. One metric that provides a concrete measurement of the added value of the service is the number of new links that the backlinking asserts into the WoD.

The added connectivity by our Backlinking service, for the case study of the UK public sector information, is depicted in the extended version of this paper¹⁷, where a table with the statistical results can be found. We complete the evaluation of the backlinking by developing a Web application¹⁸ that using the OpenSpace¹⁹ map API for the UK displays the data in a meaningful way for an end-user. The application is accessible at <http://map.psi.enacting.org> and is just one of the possible usages of our Backlinking Service.

VI. CONCLUSIONS

In this paper, we have argued that the absence of backlinks from Foreign URIs in Open Linked Data results in a navigational gap. Using Public Sector Information (PSI), we provide motivating examples as to why backlinks from Foreign URIs are especially important for items that link data from multiple publishers. These hubs provide the bridge between information resources, supplying the essential links in the WoD that will maintain a Small World network property and allow consumers to combine previously disconnected pieces of data. To alleviate this navigation problem, we present our design of a generic light-weight backlinking service and describe our implementation over UK PSI.

We have evaluated our work, demonstrating that a significant number of backlinking statements (e.g. over 1.4M statements linking statistics.data.gov.uk to education.psi.enacting.org) can be automatically generated and served to Linked Data consumers via a simple REST API. As an example of such backlinking exploitation we have developed a Web application that through a map interface is able to navigate the Web of Data displaying information from decentralized data sets.

VII. ACKNOWLEDGEMENTS

This work was supported by the EnAKTing project funded by the Engineering and Physical Sciences Research Council under contract EP/G008493/1.

¹⁷<http://eprints.ecs.soton.ac.uk/21256/>

¹⁸Refer to <http://map.psi.enacting.org/how> accessed 19/03/2010

¹⁹<http://openspace.ordnancesurvey.co.uk/openspace/> accessed 19/03/2010

REFERENCES

- [1] K. Alexander, R. Cyganiak, M. Hausenblas, and J. Zhao. Describing Linked Datasets - On the Design and Usage of void, the 'Vocabulary of Interlinked Datasets'. In *WWW 2009 Workshop: Linked Data on the Web (LDOW2009)*, Madrid, Spain, 2009.
- [2] Tim Berners-Lee. Design issues: Linked data. <http://www.w3.org/DesignIssues/LinkedData.html>, 2006.
- [3] Krishna Bharat, Andrei Broder, Monika Henzinger, and Puneet Kumar. The connectivity server: fast access to linkage information on the web. In *proceedings of the 7th International World Wide Web Conference*, 1997.
- [4] Soumen Chakrabarti, David A. Gibson, and Kevin S. McCurley. Surfing the web backwards. In *In: Proc. of WWW 8 Conference*, pages 1679–1693, 1999.
- [5] Gianluca Correndo, Manuel Salvadores, Yang Yang, Nick Gibbins, and Nigel Shadbolt. Geographical service: a compass for the web of data. In *WWW2010 Workshop: Linked Data on the Web (LDOW2010)*, April 2010.
- [6] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. Hypertext transfer protocol – http/1.1. RFC 2616 (Draft Standard), June 1999. Updated by RFC 2817.
- [7] Roy T. Fielding. *Architectural Styles and the Design of Network-based Software Architectures*. PhD thesis, University of California, Irvine, 2000.
- [8] Andrew M. Fountain, Wendy Hall, Ian Heath, and Hugh Davis. Microcosm: An open model for hypermedia with dynamic linking. In A Rizk, N Streitz, and J Andre, editors, *Hypertext: Concepts, Systems and Applications, Proceedings of ECHT'90, Paris, November 1990*, pages 298–311. Cambridge University Press, 1990.
- [9] Michael Hausenblas, Wolfgang Halb, Yves Raimond, Lee Feigenbaum, and Danny Ayers. SCOVO: Using statistics on the web of data. In *ESWC*, volume 5554 of *Lecture Notes in Computer Science*, pages 708–722. Springer, 2009.
- [10] Theodor H. Nelson. *Literary machines*. T. Nelson, 1981.
- [11] Tope Omitola, Christos Koumenides, Igor Popov, Yang Yang, Manuel Salvadores, Gianluca Correndo, Tim Berners-Lee, Nick Gibbins, Wendy Hall, mc schraefel, and Nigel Shadbolt. Put in your postcode, out comes the data: A case study. In *7th Extended Semantic Web Conference*, April May 2010.
- [12] Eyal Oren, Renaud Delbru, Michele Catasta, Richard Cyganiak, Holger Stenzhorn, and Giovanni Tummarello. Sindice.com: a document-oriented lookup index for open linked data. *IJMSO*, 3(1):37–52, 2008.
- [13] James E. Pitkow and R. Kipp Jones. Supporting the web: a distributed hyperlink database system. *Comput. Netw. ISDN Syst.*, 28(7-11):981–991, 1996.